

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 07:04:57 ; Search time 9694.27 Seconds
(without alignments)
12038.564 Million cell updates/sec

Title: US-08-731-499-12
Perfect score: 3066
Sequence: 1 GGAAACGACATGACCATGA.....GGTACCCAAATCGCCCTATA 3066

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:.*
2: gb_est2:.*
3: gb_hic:.*
4: gb_est3:.*
5: gb_est4:.*
6: gb_est5:.*
7: gb_est6:.*
8: gb_gsl1:.*
9: gb_gsl2:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 2342.4 | 76.4 | 2610 | 3 | BC024044 Homo sapi |
| 2 | 809.4 | 26.4 | 1086 | 5 | BM912146 AGENCOURT |
| 3 | 783.8 | 25.7 | 1038 | 2 | BE249939 600942990 |
| 4 | 753.2 | 24.6 | 837 | 4 | BG290366 602388204 |
| 5 | 737.8 | 24.1 | 1089 | 5 | BA19743 BX419743 |
| 6 | 696.4 | 22.7 | 897 | 7 | CN646528 ILLUMIGEN |
| 7 | 595.8 | 19.4 | 5738 | 3 | CR749643 Homo sapi |
| 8 | 595.6 | 19.4 | 2339 | 3 | AK008957 Mus muscu |
| 9 | 594.8 | 19.4 | 610 | 1 | AL120048 DKEZp761K |
| 10 | 592.6 | 19.3 | 820 | 7 | CO724689 ILLUMIGEN |
| 11 | 586.4 | 19.1 | 1076 | 4 | BG702473 602684537 |
| 12 | 583.4 | 19.0 | 598 | 5 | BP199106 BP199106 |
| 13 | 581.4 | 19.0 | 583 | 5 | BP342407 BP342407 |
| 14 | 581 | 18.9 | 589 | 4 | BM702087 UI-E-CQ1- |
| 15 | 573.2 | 18.7 | 583 | 5 | BX486692 DKEZp686P |
| 16 | 566.8 | 18.5 | 583 | 5 | BP195050 BP195050 |
| 17 | 564.2 | 18.4 | 582 | 5 | BP193074 BP193074 |
| 18 | 561.6 | 18.3 | 723 | 7 | CN647268 ILLUMIGEN |
| 19 | 561.2 | 18.3 | 578 | 5 | BP341239 BP341239 |
| 20 | 560.6 | 18.3 | 581 | 5 | BP216015 BP216015 |
| 21 | 557.4 | 18.2 | 588 | 5 | BP193198 BP193198 |
| 22 | 537.2 | 17.5 | 581 | 5 | BP326142 BP326142 |
| 23 | 536.8 | 17.5 | 677 | 4 | BP549901 603194747 |
| 24 | 532 | 17.4 | 564 | 5 | BP221049 BP221049 |

| | | | | | |
|----|-------|------|------|---|--------------------|
| 25 | 530.6 | 17.3 | 579 | 5 | BP342645 |
| 26 | 520.8 | 17.0 | 710 | 5 | BP383205 |
| 27 | 519.2 | 16.9 | 563 | 5 | BP221191 |
| 28 | 515.6 | 16.8 | 765 | 7 | CO725626 ILLUMIGEN |
| 29 | 512 | 16.7 | 582 | 5 | BP213083 |
| 30 | 510 | 16.6 | 668 | 4 | BG468930 602510955 |
| 31 | 506 | 16.5 | 676 | 7 | CK002754 AGENCOURT |
| 32 | 497.2 | 16.2 | 583 | 5 | BP195817 |
| 33 | 496.8 | 16.2 | 582 | 5 | BP308450 |
| 34 | 496 | 16.2 | 508 | 2 | BE293603 601186790 |
| 35 | 495 | 16.1 | 496 | 1 | AI573188 tr70f07.X |
| 36 | 494.2 | 16.1 | 598 | 5 | BP217439 |
| 37 | 490.6 | 16.0 | 965 | 4 | BG261180 602372912 |
| 38 | 488 | 15.9 | 588 | 5 | BP310687 |
| 39 | 486 | 15.9 | 490 | 5 | BX282664 |
| 40 | 482.4 | 15.7 | 787 | 2 | BF675198 602138126 |
| 41 | 475.8 | 15.5 | 773 | 5 | BX463597 |
| 42 | 474.8 | 15.5 | 497 | 6 | CD617341 |
| 43 | 474.8 | 15.5 | 599 | 2 | BE866183 |
| 44 | 469.2 | 15.3 | 581 | 5 | BP194681 |
| 45 | 469 | 15.3 | 1104 | 7 | CN643641 ILLUMIGEN |

ALIGNMENTS

RESULT 1
BC024044
LOCUS BC024044 2610 bp mRNA linear HTC 23-SEP-2002
DEFINITION Homo sapiens, Similar to band 83, clone IMAGE:4645428, mRNA.
ACCESSION BC024044
VERSION BC024044.1 GI:23270792
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2610)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Nataşja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 29 Row: m Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4502372
This clone has the following problem: frame shifted.
Location/Qualifiers
1. .2610
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source

```
/clone="IMAGE:4645428"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH WGC 15"
/lab_host="DH10B-R"
/note="vector: pOTB7"

ORIGIN
Query Match          76.4%; Score 2342.4; DB 3; Length 2610;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 2598; Conservative 0; Mismatches 1; Indels 235; Gaps 2;

Qy 135 GCTCCACCGACAGCCAGGCACTGGGCAGCAGCGACTGGAGACCCAGGACCCTGTGCGAGA 194
Db      |||
Qy 1 GCTCCACCGACAGCCAGGCACTGGGCAGCAGCGACTGGAGACCCAGGACCCTGTGCGAGA 60
Db      |||
Qy 195 GCAGTCCGGGTGACACGAGGGGACTGAAGATACCTCCACAGGGGCTCAGCAGGAGCAAT 254
Db      |||
Qy 61 GCAGTCCGGGTGACACGAGGGGACTGAAGATACCTCCACAGGGGCTCAGCAGGAGCAAT 120
Db      |||
Qy 255 GGGTAACCAATGAGTGTTCGCCAAAGAGTTGAAGACCAAGAGAAATGAACAGAGCAGA 314
Db      |||
Qy 121 GGGTAACCAATGAGTGTTCGCCAAAGAGTTGAAGACCAAGAGAAATGAACAGAGCAGA 180
Db      |||
Qy 315 GACTTACAGGACAAACGGCTGTGCTCTGAACGGGGTTCAGTGGTGTGTCGACCCACAC 374
Db      |||
Qy 181 GACTTACAGGACAAACGGCTGTGCTCTGAACGGGGTTCAGTGGTGTGTCGACCCACAC 240
Db      |||
Qy 375 AGTTTACAGCACTTAGAGGAAGTCGACTTGGGAATAGTGTCAAGCGGATATGTGGCCAC 434
Db      |||
Qy 241 AGTTTACAGCACTTAGAGGAAGTCGACTTGGGAATAGTGTCAAGCGGATATGTGGCCAC 300
Db      |||
Qy 435 TTCTTCCCGGAGACAAACGGAGTAAGTGTCTGCGGATGCCAACGGAAAGAAATCTTTGG 494
Db      |||
Qy 301 TTCTTCCCGGAGACAAACGGAGTAAGTGTCTGCGGATGCCAACGGAAAGAAATCTTTGG 360
Db      |||
Qy 495 GAAAGAGGCCAAACCCGAGGACCAAGCTGTCTAAATCTCGTTTCTTTTGTGATGCTCTCTCG 554
Db      |||
Qy 361 GAAAGAGGCCAAACCCGAGGACCAAGCTGTCTAAATCTCGTTTCTTTTGTGATGCTCTCTCG 420
Db      |||
Qy 555 GCCTGTACAGGACGTACCGGAGCCAGGACCAAGCGCGAGATTCATCCCTTGGATCAGTGAAGCT 614
Db      |||
Qy 421 GCCTGTACAGGACGTACCGGAGCCAGGACCAAGCGCGAGATTCATCCCTTGGATCAGTGAAGCT 480
Db      |||
Qy 615 TGATGTACGCTCCAAATAAGTCTCAGCGAAACAAAGACCCAAAGTGAAGCTGGAACATCTCC 674
Db      |||
Qy 481 TGATGTACGCTCCAAATAAGTCTCAGCGAAACAAAGACCCAAAGTGAAGCTGGAACATCTCC 540
Db      |||
Qy 675 GGTGACGCTGGAACCGGGGAGGACACAGATAAAACCCAGGGACGCCCCGGGCCAAGA 734
Db      |||
Qy 541 GGTGACGCTGGAACCGGGGAGGACACAGATAAAACCCAGGGACGCCCCGGGCCAAGA 600
Db      |||
Qy 735 CAAGTCTCTCTGCGCCAGGGATCCACGCTTCTCCACCTGAGACAGAGGGGAGCAGG 794
Db      |||
Qy 601 CAAGTCTCTCTGCGCCAGGGATCCACGCTTCTCCACCTGAGACAGAGGGGAGCAGG 660
Db      |||
Qy 795 AGGGAAGCTCCCTCCAAAGCCAAAGGACTCCAGC- TTTTTCACAAATCTTTCAGAGCTGG 853
Db      |||
Qy 661 AGGGAAGCTCCCTCCAAAGCCAAAGGACTCCAGCTTTTTCACAAATCTTTCAGAGCTGG 720
Db      |||
Qy 854 ACAAGGGAAGAAAGTGCAGGTGACGCCAAGCAGGAAGCCAAAGGCGGAGCAATC 913
Db      |||
Qy 721 ACAAGGGAAGAAAGTGCAGGTGACGCCAAGCAGGAAGCCAAAGGCGGAGCAATC 780
Db      |||
Qy 914 AAGCAAGGTGGATGAGTCTCTGCTTATCAGGGCAGTCCGATGATGCTCCCTCGAGGA 973
Db      |||
Qy 781 AAGCAAGGTGGATGAGTCTCTGCTTATCAGGGCAGTCCGATGATGCTCCCTCGAGGA 840
Db      |||
Qy 974 AGGACATAGTTGACGGCAGGAAAGAGGACAGAACTTTGGAACTCGGATTTGCTCTG 1033
Db      |||
Qy 841 AGGACATAGTTGACGGCAGGAAAGAGGACAGAACTTTGGAACTCGGATTTGCTCTG 900
Db      |||
Qy 1034 TCCCTGGGGACCCAGAAAGCACTGGAGACTGCAAGGAGCAATGCCAGGACGAGCTATAG 1093
Db      |||
```

```
Db 901 TCCCTGGGGACCCAGAAAGAGCTGGAGACTGCAAAAGSAGATTCCTCCAGGACAGACTATAG 960
Qy 1094 CAGAGATAATAATTTCCATCATGAGTTTCTTTAAACTCTGGTTTCACTTAACAAAGCTG 1153
Db      |||
Qy 961 CAGAGATAATAATTTCCATCATGAGTTTCTTTAAACTCTGGTTTCACTTAACAAAGCTG 1020
Db      |||
Qy 1154 AAACAAAAGAGCCCAAGAGACACCGGTGCTGAAAGTCAACCCACCTTCAGCTGACC 1213
Db      |||
Qy 1021 AAACAAAAGAGCCCAAGAGACACCGGTGCTGAAAGTCAACCCACCTTCAGCTGACC 1080
Db      |||
Qy 1214 TTAAGTCAGACAAAGCCAACTTTTACATCCAGGAGACCCAAAGGGGCTGGCAGAAATTTCCA 1273
Db      |||
Qy 1081 TTAAGTCAGACAAAGCCAACTTTTACATCCAGGAGACCCAAAGGGGCTGGCAGAAATTTCCA 1140
Db      |||
Qy 1274 AAGGATCGAACCCCATCGGGGCACACACAGTCCGTGACAAACCCCTGAACTCTGCGAAGGAAG 1333
Db      |||
Qy 1141 AAGGATCGAACCCCATCGGGGCACACACAGTCCGTGACAAACCCCTGAACTCTGCGAAGGAAG 1200
Db      |||
Qy 1334 GCACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAAGT 1393
Db      |||
Qy 1201 GCACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAAGT 1260
Db      |||
Qy 1394 CAGTTAAAGAGGACTCAGTCCCAACAGTGGGAGGAGAAATGCTGCTGAGTCAACCAG 1453
Db      |||
Qy 1261 CAGTTAAAGAGGACTCAGTCCCAACAGTGGGAGGAGAAATGCTGCTGAGTCAACCAG 1301
Db      |||
Qy 1454 TAGAGATTATAAAGTCCAAGGAAGTAGAATCAGCCTTTACAAACAGTGGACCTCAACGAAG 1513
Db      |||
Qy 1302 ----- 1301
Db      |||
Qy 1514 GAGATGTGCACCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAAGCAAAACCAAGAA 1573
Db      |||
Qy 1302 ----- 1301
Qy 1574 CCTCTGTATGGCGTTTCTTCAGACAAATGTTCAGTGAAGGGGATGGAGGATCACCCACT 1633
Db      |||
Qy 1302 ----- 1301
Qy 1634 CAGAAGAAATAAATGGGAAGAGACTCCAGCTGCCAAAACATCAGACTCCACAGAAAGACTA 1693
Db      |||
Qy 1302 ----- 1301
Qy 1694 TCACACCGCCAGAGCCTGAACCAAACAGGAGACACACAGAGAGGTAAAGAGGGCTCTCTCGA 1753
Db      |||
Qy 1327 TCACACCGCCAGAGCCTGAACCAAACAGGAGCACACACAGAGGGTAAAGAGGGCTCTCTCGA 1386
Db      |||
Qy 1754 AGGACAAAGAGTCAAGCAGCCGAGATGAACAAGCAGAGAGAGCAACAAGCAGAGAGCCAAAG 1813
Db      |||
Qy 1387 AGGACAAAGAGTCAAGCAGCCGAGATGAACAAGCAGAGAGAGCAACAAGCAGAGAGCCAAAG 1446
Db      |||
Qy 1814 AACCCAGCCAGTGCACAGAGCAGGCCACCGTGGACACGAACTCACTGCAGAAATGGGGACA 1873
Db      |||
Qy 1447 AACCCAGCCAGTGCACAGAGCAGGCCACCGTGGACACGAACTCACTGCAGAAATGGGGACA 1506
Db      |||
Qy 1874 AGCTCCAAAAGAGACCTTGAGAAAGCGGACAGTCCCTTTGGGGGCTTCTTTAAAGGCTGG 1933
Db      |||
Qy 1507 AGCTCCAAAAGAGACCTTGAGAAAGCGGACAGTCCCTTTGGGGGCTTCTTTAAAGGCTGG 1566
Db      |||
Qy 1934 GACCAAAAGCGGATGTTGGATGCTCAAGTGCNAAACAGACCCAGTATCCATCGGACCAAGTTG 1993
Db      |||
Qy 1567 GACCAAAAGCGGATGTTGGATGCTCAAGTGCNAAACAGACCCAGTATCCATCGGACCAAGTTG 1626
Db      |||
Qy 1994 GCAAAACCAAGTAAACAAATCAGCACGCTTCCACACAGGTTCTCTCTGCCACCAAGATGTG 2053
Db      |||
Qy 1627 GCAAAATCCAGTAAACAAATCAGCACGCTTCCACACAGGTTCTCTCTGCCACCAAGATGTG 1686
Db      |||
Qy 2054 TTCTCTTACTTCCATCTCTCTCCCAACACGCTCCATGTATATATTTCTTCTGTATGGCCAG 2113
Db      |||
Qy 1687 TTCTCTTACTTCCATCTCTCTCCCAACACGCTCCATGTATATATTTCTTCTGTATGGCCAG 1746
Db      |||
Qy 2114 CAAATGAATTTCTGCTAGAAATTAAGCCGAGCTGTTGTATATTTGAGGTGTATTTATTTA 2173
Db      |||
Qy 1747 CAAATGAATTTCTGCTAGAAATTAAGCCGAGCTGTTGTATATTTGAGGTGTATTTATTTA 1806
Db      |||
```

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLCM1985 row: j column: 12
High quality sequence start: 13
High quality sequence stop: 510.
Location/Qualifiers
1. .1086
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5473643"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

JOURNAL COMMENT

FEATURES source

ORIGIN

Query Match 26.4%; Score 809.4; DB 5; Length 1086;
Best Local Similarity 95.4%; Pred. No. 3.3e-121;
Matches 867; Conservative 0; Mismatches 36; Indels 6; Gaps 3;

QY 2174 CGTCTCGTCCAGTCTTTCTGGCAATAACAGTAAAGATGGTTTACGAGTCACTAG 2233
DB 1807 CGTCTCGTCCAGTCTTTCTGGCAATAACAGTAAAGATGGTTTACGAGTCACTAG 1866
QY 2234 TTGGGTGAGAGAGTGCATGATCACCAGCAGGAAAGGAGGAAATAGAGGAATGTGTT 2293
DB 1867 TTGGGTGAGAGAGTGCATGATCACCAGCAGGAAAGGAGGAAATAGAGGAATGTGTT 1926
QY 2294 GGGTTAAGTGAATAATGGCAGTGGTGGCGGGCGTGGTGGCTCCGCTGTAAATCTCA 2353
DB 1927 GGGTTAAGTGAATAATGGCAGTGGTGGCGGGCGTGGTGGCTCCGCTGTAAATCTCA 1986
QY 2354 GCATTTGGGAGGCGGAGGCGAGTGGATCACTGAGGTCAGAGTTCAAGACTAGCCTGG 2413
DB 1987 GCATTTGGGAGGCGGAGGCGAGTGGATCACTGAGGTCAGAGTTCAAGACTAGCCTGG 2046
QY 2414 CCAACATCATGAAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGGCATGGTGGCAC 2473
DB 2047 CCAACATCATGAAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGGCATGGTGGCAC 2106
QY 2474 ACACCTGTAGTCCAGTACTCGGAGCCCAACGACAGAAACCGCTTTGTACCCAGGAGG 2533
DB 2107 ACACCTGTAGTCCAGTACTCGGAGCCCAACGACAGAAACCGCTTTGTACCCAGGAGG 2166
QY 2534 TGGAGGTTGCAGTGAAGCCGAAGTTGCACATTCGACTCCACCTGGGCGACAGAGCAAGA 2593
DB 2167 TGGAGGTTGCAGTGAAGCCGAAGTTGCACATTCGACTCCACCTGGGCGACAGAGCAAGA 2226
QY 2594 TTCTATCAAAAAAAGGCGAGTGGCAAGTAAATATAGAGAGAAATGCTGTAGAGG 2653
DB 2227 TTCTATCAAAAAAAGGCGAGTGGCAAGTAAATATAGAGAGAAATGCTGTAGAGG 2286
QY 2654 AATTAAGCGTTGTAGTAAACGCGTCTCATCTCTAAGCTTTGAAGAGGAGACGAAAT 2713
DB 2287 AATTAAGCGTTGTAGTAAACGCGTCTCATCTCTAAGCTTTGAAGAGGAGACGAAAT 2346
QY 2714 CCATTTGTTTAAATTCACATCTCAAGGAGGAGAAACCGGGGTGTGTGGGTGTGCTCA 2773
DB 2347 CCATTTGTTTAAATTCACATCTCAAGGAGGAGAAACCGGGGTGTGTGGGTGTGCTCA 2406
QY 2774 ATTTCTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTCAA 2833
DB 2407 ATTTCTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTCAA 2466
QY 2834 ATAGGTCCTTTGTAAGTTATTGATGAGAGGAGAAAGATTGACTGGGAGGGCTTAAATG 2893
DB 2467 ATAGGTCCTTTGTAAGTTATTGATGAGAGGAGAAAGATTGACTGGGAGGGCTTAAATG 2526
QY 2894 ATTTGGGAAAAACAATGCTTTTGGGCTCAGTGACACCGGCAAGATTACAACTTAAAAA 2953
DB 2527 ATTTGGGAAAAACAATGCTTTTGGGCTCAGTGACACCGGCAAGATTACAACTTAAAAA 2586
QY 2954 AAAAAAATAAAAAA 2967
DB 2587 AAAAAAATAAAAAA 2600

RESULT 2
BM912146
LOCUS BM912146
DEFINITION AGENCOURT_6613041 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5473643
5', mRNA sequence.
ACCESSION BM912146
VERSION BM912146.1 GI:19362525
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1086)
NIH-MGC <http://mgi.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

```
Db 541 GGACTCCAGCTTTTGTGACAAATCTTCAAGCTGGACAAAGGCACAGGAAAGGTGCCAGG 600
Qy 879 TGACAGCCAAACAGGAAGCCAGAGGGCAGAGCATCAAGACAAAGTGGATGAGTTCCTGG 938
Db 601 TGACAGCCAAACAGGAAGCCAGAGGGCAGAGCATCAAGACAAAGTGGATGAGTTCCTGG 660
Qy 939 CTTATCAGGGGAGTCCGATGATGTCCCTGACAGGGAAGGACATAGTTGACGGCAAGGAAA 998
Db 661 CTTATCAGGGGAGTCCGATGATGTCCCTGACAGGGAAGGACATAGTTGACGGCCACGAGA 720
Qy 999 AGAAGGACAAAGCTTGAACTGGCGATGCTCTGCTCGGAGACCCAGAAAGACTGGA 1058
Db 721 CGAAGGACCAAGAACTTGCACTGGCGATGCTCTGCTCGGAGACCCAGAAAGACTGGA 780
Qy 1059 GACTGCAAGGACGATTCACAGGAGCAGCTATAGCAGAG--AATAAATAATTCATCATGA 1117
Db 781 GACTGCCAAGGACGATTCACAGGAGGAGTATAGCCAGAAATAATAATTTCCATGCTGA 840
Qy 1118 GTTTCCTTTT--AAAACTCTGGTTTTCACCTTAACAAAGCT---GAAACAAAAGGACCCAGA 1172
Db 841 GTTTCCTTTTAAACTCTGGTTTTCACCTTACAAAGCTTGAGAACCAAGAAAGACCCGAA 900
Qy 1173 AGACACGGG 1181
Db 901 AAAAACCGG 909
```

```
RESULT 3
BE249939
LOCUS BE249939F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:2959439 5',
DEFINITION BE249939 1038 bp mRNA linear EST 13-JUL-2000
ACCESSION BE249939
VERSION BE249939.1 GI:9120042
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1038)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
```

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW51 row: c column: 24
High quality sequence stop: 784.

```
FEATURES
source
1..1038
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2959439"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_15"
/note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
```

ORIGIN

```
Query Match 25.7%; Score 788.8; DB 2; Length 1038;
Best Local Similarity 94.5%; Pred. No. 7,2e-118;
Matches 862; Conservative 0; Mismatches 42; Indels 8; Gaps 4;
Qy 656 GTGAGAGCTGGACACTTCCGGTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCAG 715
Db 1 GTGAGAGCTGGACACTTCCGGTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCAG 60
Qy 716 GGACAGCCCCGGCCAAAGACAAGGTCTCTCTGCGCCAGGGATCCCAACGCTTCTCCAC 775
Db 61 GGACAGCCCCGGCCAAAGACAAGGTCTCTCTGCGCCAGGGATCCCAACGCTTCTCCAC 120
Qy 776 CTGAGACAGGGGGAGCAGGAGGAGAGCTCCCTCCAAGCCCAAGAGTCTCAGCTTTTGTG 835
Db 121 CTGAGACAGGGGGAGCAGGAGGAGAGCTCCCTCCAAGCCCAAGAGTCTCAGCTTTTGTG 180
Qy 836 ACAATTTCTTCAAGCTGGACAAAGGACAGGAAAGGTGCCAGGTGACAGCCCAAGAG 895
Db 181 ACAATTTCTTCAAGCTGGACAAAGGACAGGAAAGGTGCCAGGTGACAGCCCAAGAG 240
Qy 896 CCAAGAGGGCAGAGCATCAAGACAAGGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCG 955
Db 241 CCAAGAGGGCAGAGCATCAAGACAAGGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCG 300
Qy 956 ATGATGTCCCTGACAGGGAAGACATAGTTGACGGCAAGGAAAAAGAGGACAAAGACTTG 1015
Db 301 ATGATGTCCCTGACAGGGAAGACATAGTTGACGGCAAGGAAAAAGAGGACAAAGACTTG 360
Qy 1016 GAACCTGGGATTGCTCTGCTCCCTGGGACCCAGAGAGACTGGAGACTGCCAAGGACGAT 1075
Db 361 GAACCTGGGATTGCTCTGCTCCCTGGGACCCAGAGAGACTGGAGACTGCCAAGGACGAT 420
Qy 1076 CCCAGGACAGCAGCTATAGCAGAGAAATAATAATTCATCATGAGTTCCTTTAAAACTCTGG 1135
Db 421 CCCAGGACAGCAGCTATAGCAGAGAAATAATAATTCATCATGAGTTCCTTTAAAACTCTGG 480
Qy 1136 TTTCACTTAAACAAAGCTGAAACAAAAAGGACCCAGAGACACCGGTGCTGAAAGTCAAC 1195
Db 481 TTTCACTTAAACAAAGCTGAAACAAAAAGGACCCAGAGACACCGGTGCTGAAAGTCAAC 540
Qy 1196 CCACACTTCAGCTGACCTTAAAGTCAGACAAAGCCAACTTTACATCCAGAGAGACCCAG 1255
Db 541 CCACACTTCAGCTGACCTTAAAGTCAGACAAAGCCAACTTTACATCCAGAGAGACCCAG 600
Qy 1256 GGGCTGCAAGAAATTCACAAAGGATGCAACCATCCGGGCACACACAGTCCCTGTGACAAACC 1315
Db 601 GGGCTGCAAGAAATTCACAAAGGATGCAACCATCCGGGCACACACAGTCCCTGTGACAAACC 660
Qy 1316 CTGAACCTGCAAGGAAAGGACCAAGGAGAAATCAGGACCCACCTCTCTCTCTGCTCTGGGCA 1375
Db 661 CTGAACCTGCAAGGAAAGGACCAAGGAGAAATCAGGACCCACCTCTCTCTCTGCTCTGGGCA 720
Qy 1376 AACTGTTTGGAAAAAGTCAGTTAAAGAGNACTCAGTCCCCACACAGGTG--CGGAGGAGAT 1434
Db 721 AACTGTTTGGAAAAAGTCAGTTAAAGAGNACTCAGTCCCCACACAGGTG--CGGAGGAGAT 780
Qy 1435 GTGGTG--TGTCAGTCAACAGTAGAGATTATAAGTCCAAAGGAAGTAGAATCAGGCTTACA 1493
Db 781 GTGGTGTTGTGAGTCCCAAGTAGAGAT--TTAAGTCCCGCCAGTTGATCTCGCTACAA 838
Qy 1494 AACAGTGGACCTCAACGAAGGAGATGCTGCACTGAAACCCACAGAGGAAACTCAAAAG 1553
Db 839 CCAGGGGGCACTACCCGAGGA---ATGCGGCTCTGAACCCCAAGAGGCAATCTCAAGAG 894
Qy 1554 AGAAGAAAGCA 1565
Db 895 AAAAGGCAACCA 906
RESULT 4
BE290366
LOCUS BE290366F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4517041 5',
DEFINITION
```


Best Local Similarity 80.1%; Pred. No. 1.3e-109;
Matches 1013; Conservative 8; Mismatches 6; Indels 237; Gaps 4;

Qy 751 GCCAGGATCCAC-GCTTCTCCACCTGACAGAGGGGAGCAGAGAGAGCTCCCTC 809
Db 1 GCCAGGATCCACAGAGCTTCTCCACCTGACAGAGGGGAGCAGAGAGAGCTCCCTC 60

Qy 810 CAAGCCAGAGCTCCAGCTTTTGGCAAAATCTTCAAGCTGACAGGAGGACAGGAAA 869
Db 61 CAAGCCAGAGCTCCAGCTTTTGGCAAAATCTTCAAGCTGACAGGAGGACAGGAAA 120

Qy 870 GGTGCCAGGTGACAGCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 929
Db 121 GGTGCCAGGTGACAGCCAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180

Qy 930 GGTCTCGCTTATCAGGGAGTCCGATGATGTCCTGAGGAGGAGGAGGAGGAGGAGG 989
Db 181 GGTCTCGCTTATCAGGGAGTCCGATGATGTCCTGAGGAGGAGGAGGAGGAGGAGG 240

Qy 990 CAAGGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1049
Db 241 CAAGGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

Qy 1050 AGGACTGAGACTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1109
Db 301 AGGACTGAGACTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360

Qy 1110 CATCATGAGTTCTTAAACCTGTTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 1169
Db 361 CATCATGAGTTCTTAAACCTGTTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 420

Qy 1170 AGAAGACACGGGTGCTGAAAAGTCAACCACTTCAAGCTGAGTCAAGCTGAGTCAAGCTGAGT 1229
Db 421 AGAAGACACGGGTGCTGAAAAGTCAACCACTTCAAGCTGAGTCAAGCTGAGTCAAGCTGAGT 480

Qy 1230 CAACCTTATCATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1289
Db 481 CAACCTTATCATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540

Qy 1290 GGGGACACACAGTCCGTGACAAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGA 1348
Db 541 GGGGACACACAGTCCGTGACAAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCTTGA 600

Qy 1349 CAGGACCCACCTCTGCTCTGCGCAAACTGTTTGGAAAAGTCAAGTAAAGAGGACT 1408
Db 601 CAGGACCCACCTCTGCTCTGCGCAAACTGTTTGGAAAAGTCAAGTAAAGAGGACT 660

Qy 1409 CAGTCCCCACAGGTGCGGAGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1468
Db 661 CAGTCCCCACAGGTGCGGAGGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 686

Qy 1469 CCAAGGAGTGAATCAGCCTTACAAACAGTGGACCTCAACAGGAGGAGTGTGTGTGTGTGTGTGT 1528
Db 687 ----- 686

Qy 1529 AACCCAGAGCGAACTCAAAAG 1588
Db 687 ----- 686

Qy 1589 TTCTCAGAAATGTCAAGTGAAGGGGATGAGGGATCACCCTCACTCAGAGAGAGAGAGAGAGAG 1648
Db 687 ----- 686

Qy 1649 GGAAGACTCCAGTGCCCAACATCAGACTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1708
Db 687 -----ACATCAGACTCCACAG 726

Qy 1709 CTGACCAACAGGAGCACAACAGAGGAGTAAAGAGGCTCTCGAAGGAGAGAGAGAGAGAGAGAG 1768
Db 727 CTGACCAACAGGAGCACAACAGAGGAGTAAAGAGGCTCTCGAAGGAGAGAGAGAGAGAGAGAG 786

Qy 1769 CAGCCGAGTGAACAG 1828
Db 1769 CAGCCGAGTGAACAG

Db 787 CAGCCGAGATGAACAGCAGAGAGAGCAACAGCAGAGAGCAAAAGAACCCAGCCAGTGCA 846

Qy 1829 CAGAGCAGGCCACGGTGGACACGAACCTCACTGCAGATGGGACCAAGCTCCAAAGAGAC 1888

Db 847 CAGAGCAGGCCACGGTGGACACGAACCTCACTGCAGATGGGACCAAGCTCCAAAGAGAC 906

Qy 1889 CTGAGAGCGGCAGCAGCTCCCTTTGGGGCTTCTTTT-AAAGCCCTGGGACCAAAAGCGGATG 1947

Db 907 CTGAGAGCGGCAGCAGCTCCCTTTGGGGCTTCTTTAAAGGCTTGGGACCAAAAGCGGATG 966

Qy 1948 TTGATGCTCAAGTGAACAGAGCCAGATGTCATCGGACAGCTTGGCAAAACCAAGTAA 2007

Db 967 TTGATGCTCAAGTGAACAGAGCCAGATGTCATCGGACAGCTTGGCAAAATCCAAARWAA 1026

Qy 2008 ACBA 2011

Db 1027 AMAA 1030

RESULT 6

CN646528

LOCUS

DEFINITION CN646528 897 bp mRNA linear EST 13-MAY-2004

ILLUMIGEN MQ0 26475 Katze_MBR Macaca mulatta cDNA clone IBIUW.8476

5' similar to bases 146 to 797 highly similar to human BCAS1

(Hs.400556), mRNA sequence.

ACCESSION CN646528

VERSION CN646528.1 GI:47159971

KEYWORDS EST.

SOURCE Macaca mulatta (rhesus monkey)

ORGANISM Macaca mulatta

REFERENCE 1 (bases 1 to 897)

AUTHORS Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magnus,C.L.

TITLE Large-scale Rhesus Macaque cDNA Sequencing

JOURNAL Unpublished (2003)

COMMENT Contact: C. Magnus

Illumigen Biosciences Inc.

2203 Airport Way S, Suite 450, Seattle, WA 98134, USA

Tel: 2063780400

Fax: 2063780408

Email: cmagness@illumigen.com

Sequenced on 2004.03.10. 746 Q20 bases. Contains sequence similar to Alu

PCR Primers

FORWARD: CCCTCACTAAAGGGAACAAA

BACKWARD: CACTATAGGCGAATGGGTA

Insert Length: 897 Std Error: 0.00

Plate: CL000144 row: G column: 09

Seq primer: CCCTCACTAAAGGGAACAAA

POLYA=No.

FEATURES

Location/Qualifiers

1..897

/organism="Macaca mulatta"

/mol_type="mRNA"

/strain="Indian"

/db_xref="taxon:9544"

/clone="IBIUW.8476"

/sex="female"

/dev_stage="adult"

/lab_host="E. coli SOLR"

/clone_lib="katze_MBR"

/note="Organ: brain; Vector: Uni-ZAP XR; Site 1: EcoR I; Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis kit (catalog #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit (catalog #200450)"

ORIGIN

Query Match 22.7%; Score 696.4; DB 7; Length 897;

Best Local Similarity 95.2%; Pred. No. 6.7e-103; Indels 0; Gaps 0;

Matches 718; Conservative 0; Mismatches 36;

```
Qy 1667 AAACATCAGACTCCACAGAAAGACTATCACACGCCAGAGCCTGAACCAACAGGAGCAC 1726
Db 144 AGACATCAGACTCCAGAGAAAGACTATCACACACAGAGCCTGAACCAACAGGAGGC 203
Qy 1727 CACAGAGGGTAAAGAGGGCTCTTCCAGAGGACCAAGAGTCAAGAGCAGAGATGAACAAGC 1786
Db 204 CCCAGAGGGCAAGAGGGCTCTTCCAGAGGACCAAGAGTCAAGAGCAGAGATGAACAAGC 263
Qy 1787 AGAGAGCAACAGCAGGAGGACCAAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGC 1846
Db 264 AGAGAGCAACAGCAGGAGGACCAAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGC 323
Qy 1847 ACACGAATCTACTGACAGATGGGACCAAGTCTCAAAAGAGAGCCTCAGAGAGGCGCAGCAGT 1906
Db 324 ACACGAATCTACTGACAGATGGGACCAAGTCTCAAAAGAGAGCCTCAGAGAGGCGCAGCAGT 383
Qy 1907 CCCTGGGGGCTCTTTTAAAGGCTGGGACCAAGTCTCAAAAGAGAGCCTCAGAGTGAACA 1966
Db 384 CCCTGGGGGCTCTTTTAAAGGCTGGGACCAAGTCTCAAAAGAGAGCCTCAGAGTGAACA 443
Qy 1967 CAGACCCAGTATCCATCGACAGTGGGACCAAGTCTCAAAAGAGAGCCTCAGAGTGAACA 2026
Db 444 CAGACCCAGTATCCATCGACAGTGGGACCAAGTCTCAAAAGAGAGCCTCAGAGTGAACA 503
Qy 2027 ACCAGGTTCTCTGACCAAGAGTGTCTCTTACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 2086
Db 504 ACCAGGTTCTCTGACCAAGAGTGTCTCTTACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 563
Qy 2087 CCATGTATATATCTTCTGATGGGACCAAGTGAATTTCTGCTAGAAATTAAGCCCGAG 2146
Db 564 CCATGTATATATCTTCTGATGGGACCAAGTGAATTTCTGCTAGAAATTAAGCCCGAG 623
Qy 2147 CTGTTGTATATAGGTGTATTTTATAGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 2206
Db 624 CTGTTGTATATAGGTGTATTTTATAGTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 683
Qy 2207 GTAAAGATGGTTTACAGGTCACCTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2266
Db 684 GTAAAGATGGTTTACAGGTCACCTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 743
Qy 2267 AAAGGGAGGAGTAAAGGAATGTGTTCCGGTAAAGTGAATGAATGAATGAATGAATGAATGA 2326
Db 744 AAAGGGAGGAGTAAAGGAATGTGTTCCAGTTTAAAGTGAATGAATGAATGAATGAATGAAT 803
Qy 2327 GCGTGGTGGTCTGCGCTGTATCTCAGCATTGTTGGAGGCGGAGGAGGAGGAGGAGGAGGAG 2386
Db 804 GTGTTGGTGGTCTGCGCTGTATCTCAGCATTGTTGGAGGCGGAGGAGGAGGAGGAGGAGG 863
Qy 2387 GAGTCCAGGAGTCAAGACTAGCCTGGCCACAT 2420
Db 864 GAGTCCAGGAGTCAAGACTAGCCTGGCCACAT 897

RESULT 7
CR749643 5738 bp mRNA linear HTC 19-AUG-2004
LOCUS Homo sapiens mRNA; cDNA DKFP686B1850 (from clone DKFP686B1850).
DEFINITION Homo sapiens mRNA; cDNA DKFP686B1850 (from clone DKFP686B1850).
ACCESSION CR749643
VERSION CR749643.1 GI:51476981
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5738)
Ottewaelder, B., Obermaier, B., Deutschenbauer, S., Schaiipp, A.,
Wiemann, S., Weill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
```

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKFP686B1850) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneId=DKFP686B1850
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

source

1..5738
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP686B1850"
/tissue_type="amygdala"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="breast carcinoma amplified sequence 1, N-terminus
truncated, differentially spliced"
1..5738
/gene="DKFP686B1850"
-3..1025
/gene="DKFP686B1850"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18437.1"
/db_xref="GI:51476982"
/translation="SOAATAENNNNSIMSPKTLVSPNKAKTKDPEDTASKABSVCD
GQAGKTSIQARCTGKKHLDSPRLGLAFKPKFRHKGAEKPTTSADLKSANFTSQ
ETQAGKNSGKNSGHTSVTTPPAKEDTKESGPTSLPLGLFKWKKVVCESVEI
IKSEVSALQTVLDNEGDAPEPEAKLRBEESKPTSLMAFLQMTSDTEKTIPT
PEPTGAPQKGGSSKKSAEMNKKSNKQEAKEPAQCTQATQVDTNSLQNGDK
LQKPEKRRQSLGGFFKGLGPKRLMDAQVQDTPVISGVPKPKVTGKCVFHVKKKWP
FQEWSY"

gene

CDS

19.4%; Score 595.8; DB 3; Length 5738;
Best Local Similarity 85.4%; Pred. No. 8.4e-87;
Matches 734; Conservative 0; Mismatches 17; Indels 108; Gaps 2;
Qy 1175 ACACGGTGTGAAAAGTCAACCCACCTTACAGTGCAGTCAAGAGGCAACT 1234
Db 238 ATAAAGGTGTGAAAAGTCAACCCACCTTACAGTGCAGTCAAGAGGCAACT 297
Qy 1235 TTACATCCAGGAGACCAAGGGCTGCGAAGAAATTCAGAGGATGCAACCCATCGGGC 1294
Db 298 TTACATCCAGGAGACCAAGGGCTGCGAAGAAATTCAGAGGATGCAACCCATCGGGC 357
Qy 1295 ACACAGTCCGTGACAAACCCCTGAACTGCGAAGGAGGACCAAGAGAGAAATCAGGAC 1354
Db 358 ACACAGTCCGTGACAAACCCCTGAACTGCGAAGGAGGACCAAGAGAGAAATCAGGAC 417
Qy 1355 CCACCTCTGCTCTGGGCAACTGTTTGGAAAAAGTCAAGTAAAGAGACTCAGTCC 1414
Db 418 CCACCTCTGCTCTGGGCAACTGTTTGGAAAAAGTCAAGTAAAGAGACTCAGTCC 455
Qy 1415 CCACAGTCCGGAGGAGNATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1474
Db 456 -----GT 495
Qy 1475 AAGTAGAATCAGCTTACAAACAGTGGACCTCAACAGGAGGAGATGCTGCACCTGAACCCA 1534
Db 496 AAGTAGAATCAGCTTACAAACAGTGGACCTCAACAGGAGGAGATGCTGCACCTGAACCCA 555
Qy 1535 CAGAGGCAACTCAAAAGAGAGAGAAACCAAGAACTCTCTGTAGTGGCTTTCTCA 1594
Db 556 CAGAGGCAACTCAAAAGAGAGAGAAACCAAGAACTCTCTGTAGTGGCTTTCTCA 615
Qy 1595 GACAAATGTCAAGTAAAGGGGATGGGGATCACCCTCAGAGNAAATTAATGGGAAG 1654
Db 616 GACAAATG----- 623

ORIGIN

Query Match 19.4%; Score 595.8; DB 3; Length 5738;
Best Local Similarity 85.4%; Pred. No. 8.4e-87;
Matches 734; Conservative 0; Mismatches 17; Indels 108; Gaps 2;
Qy 1175 ACACGGTGTGAAAAGTCAACCCACCTTACAGTGCAGTCAAGAGGCAACT 1234
Db 238 ATAAAGGTGTGAAAAGTCAACCCACCTTACAGTGCAGTCAAGAGGCAACT 297
Qy 1235 TTACATCCAGGAGACCAAGGGCTGCGAAGAAATTCAGAGGATGCAACCCATCGGGC 1294
Db 298 TTACATCCAGGAGACCAAGGGCTGCGAAGAAATTCAGAGGATGCAACCCATCGGGC 357
Qy 1295 ACACAGTCCGTGACAAACCCCTGAACTGCGAAGGAGGACCAAGAGAGAAATCAGGAC 1354
Db 358 ACACAGTCCGTGACAAACCCCTGAACTGCGAAGGAGGACCAAGAGAGAAATCAGGAC 417
Qy 1355 CCACCTCTGCTCTGGGCAACTGTTTGGAAAAAGTCAAGTAAAGAGACTCAGTCC 1414
Db 418 CCACCTCTGCTCTGGGCAACTGTTTGGAAAAAGTCAAGTAAAGAGACTCAGTCC 455
Qy 1415 CCACAGTCCGGAGGAGNATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1474
Db 456 -----GT 495
Qy 1475 AAGTAGAATCAGCTTACAAACAGTGGACCTCAACAGGAGGAGATGCTGCACCTGAACCCA 1534
Db 496 AAGTAGAATCAGCTTACAAACAGTGGACCTCAACAGGAGGAGATGCTGCACCTGAACCCA 555
Qy 1535 CAGAGGCAACTCAAAAGAGAGAGAAACCAAGAACTCTCTGTAGTGGCTTTCTCA 1594
Db 556 CAGAGGCAACTCAAAAGAGAGAGAAACCAAGAACTCTCTGTAGTGGCTTTCTCA 615
Qy 1595 GACAAATGTCAAGTAAAGGGGATGGGGATCACCCTCAGAGNAAATTAATGGGAAG 1654
Db 616 GACAAATG----- 623

SEGI PRSBNKSSCOTSNSVEKTPSPPEPAGTAQKNKETSSSKDKKSVKKS
TENSCKNGKQREVRPAFCVOPPTVEANAMTGDKTPKSEKRQSLGFLKGLGPR
MSDAQVQDTPVSGIPVGHKSK

polyA_signal

/note="putative"

polyA_site

/note="putative"

ORIGIN

Query Match 19.4%; Score 595.6; DB 3; Length 2339;

Best Local Similarity 64.8%; Pred. No. 1.1e-86;

Matches 1095; Conservative 0; Mismatches 414; Indels 181; Gaps 7;

```
OY 796 GGAGAGCTCCCTCCCAAGCCCAAGACTCCAGCTTTTGTGACAAATCTTCAAGCTGCAC 855
DB 55 GCAGAGCCCGGCCGACGAGGATTTGCGCTTTTGAACAGATCTTTAAACTGGAC 114
OY 856 AAGGACAGGAAAGGTGCCAGGTGCACAGCCAAAC---AGGAAGCCAAAGAGGCGCAGAGCAT 912
DB 115 AAAGGACAGGAAAGCGCGCGGTGAAACAGCCAGCCCAAGAAAGGAGGCTCGGAGAC 174
OY 913 CAAGACAGGTGGATGAGGTTCTCGCTTATACAGGCGAGTCCGATGATGTCCTCGAGGG 972
DB 175 CCAGAACAGGCCACAGAGGCTCTCGCGTGCAGGGAATCCCATGTTGTTCTTCGAGGG 234
OY 973 AAGGACATGTTGACGCAAGGAAAGAAAGAGGACAAAGAACTTGGAACTGCGGATGCTCT 1032
DB 235 GAGGACATAGTCGACAGCGAGGAG---GAGGACAAAGCGTTGACACTCTGAGTTATTCT 291
OY 1033 GTCCCTGGGACCCCAAGAGGACTGGAGACTGCMAAGGACGATTCGCCAGGACAGCTATA 1092
DB 292 GTCCCTGGGATCCAGAGTGCCGGGACACGAGAGGAGGCCCGCAGGTGTCGATACC 351
OY 1093 GCAGAGATAATAATTCATCATGATGTTCTTTAAACTCTGTTTACCTAAACAAGCT 1152
DB 352 ACAGAGACAGCAGCTCCATCATGAGCTTCTTCAAGACACTGGTTTCACTTAAACAAGCT 411
OY 1153 GAAACAAAGGACCCAGAGACAC-----1178
DB 412 GAAACAAAGGACCCAGAGACACAGGAAGCAACCAAGGACAGAGTGTCTGTGATGA 471
OY 1179 -----1178
DB 472 CACGCTGTCGCCAAGATGTCGAGAGCAGGCTAAAGCAAGAGAGGCGCTGGAC 531
OY 1179 -----GGGTGCTGAAAGTCA 1194
DB 532 AGCCCGAGGTAGGACTCTCTTTTAAAGCTCTTTAGGCACAAAGGATACTGAGAACTCA 591
OY 1195 CCCACCACTTCACTGACCTTAAGTCAGACAAAGCCAACTTTACATCCAGGAGACCCAA 1254
DB 592 CCCACGACTTCAGCCAAATCTCAAGTCGACAAAGCCAACTTTCACACCCCGAGAGACCGA 651
OY 1255 GGGGCTGGCAAGAAATTCACAAAGGATGCAACCCATCGGGGCACACACAGCTCGTGACACC 1314
DB 652 GGAAGACAAAGGCTACCAAAAGCTGACAGCCCCCACCAGCCACTC-----CTGAACCC 705
OY 1315 CTTGAACCTCGAAGGAGGACCAAGGAGAGAAATCAGGACCCACCTCTCTGCTCTGGGC 1374
DB 706 ACCAGCAAGGAGAGACAGCGCAAGGAGAGAGGCGAGGCCACCTCACTACCGCTGGGA 765
OY 1375 AAACGTTTTGGAAGGATGAGTTAAAGAGACTCAGTCCCAACAGGTGCGGAGAGAGAT 1434
DB 766 AAGTTGTTTTGGAAGAGTCAAGTTAAGGAGGATACACTTCCACAGGTGTCAGAGAGAAC 825
OY 1435 GTGTGTGTGAGTCAACAGTACAGATTAATTAAGTCCRAGGAGTAGATCAGCTTACAA 1494
DB 826 GCGGTGTGTGAGTCAACAGTACAGACCTGATAGGCTTTGAGGAGTAGATTCAGCTTACA 885
OY 1495 ACAGTGGACCTCAACAGAGGAGAGTGTGCACCTGAAACCCACAGAGGCGAAACTCAAGA 1554
DB 886 ACTGTGATCTCAGTGA---AGAGACCCAGCTGACCCACAGAGCTAAAGTCAAGAA 942
```

```
OY 1555 GAAGAAACCAACCAAGAACCTCTGATGGGTTTCTCAGACAAATGTCACTGAAAGG 1614
DB 943 GAAAGCAACCCCGGAGACCCCTCTGATGGGTTTCTCAGACAAATGTCACTGAGATCG 1002
OY 1615 GATGAGGAGTACCCACTCAGAAAGAAATAAATGGAAAGACTCCAGCTGCCAAACATCA 1674
DB 1003 AGCGAAGGATCCCGCTCGGAGAAAGTAACTGTAAGACTCCAGCTGCCAAACGTC 1062
OY 1675 GACTCCACAGAAAGACTATCACCAGCCAGAGCTTGAAACCAAGGAGACCAAGAG 1734
DB 1063 AACTCCGTGGAAGAGACGCCCTCACCAGAGCCAGAGCCCGCGGAAACAGCTCAGAAA 1122
OY 1735 GGTAAAGAGGGCTCC-----TCGAAGGACAAAGCTGACGACCCGAG 1776
DB 1123 AACAGGAGACTCTCTCTCGAAGGACAAAGAAATCAGTGGACAAAGTCACTGAGCTGAG 1182
OY 1777 ATGAACAGCAGAGAGCAACAGCAGGAGAAAGCCAAAGCCAGCTGCACAGACGAG 1836
DB 1183 AACAGCAGCAGAGAGACGCGCAACAGAGACTCAGAGAGCCGCGCCGTGTGTGACGCG 1242
OY 1837 GCCACGCTGGACACGAACTCACTGAGAAATGGGACAAAGCTCCAAAGAGACCTCGAAG 1896
DB 1243 CCCACAGTGGAGGCAACGCAATGCAGACTGGGACAAAGACCCCAAGAAAGTCTGAGAG 1302
OY 1897 CGCAGCAGTCCCTTGGGGCTTCTTTAAAGCCCTGGGACCAAGCCGATGTTGATGCT 1956
DB 1303 CGGCGACAGTCCCTCGGGGGCTTCTTGAAGGGCTTGGGACCAAGCCGATGTCGATGCT 1362
OY 1957 CAAGTGCAACAGACCCAGATATCCATCGGACAGTTGGCAACCCCAAGTAAACAAATCAG 2016
DB 1363 CAGGTGCAACAGGACCCGCTCTCCATCGGACAGTTGGAAATCCAGTAAACAAATCA 1422
OY 2017 CACGGTCCCAACAGGTTCTCTGCCACCAAGATGTGTTCTCTTCTTCTCTCTCTCT 2076
DB 1423 CCCAGTTCCTCACTGGGTTCTCTCCAC-----TCCATCTACTGTCCAGACCCAC 1472
OY 2077 CAACACGCTCCATGATATATCTCTGATGGCCAGCAATGAAATTCGCTAGAAAT 2136
DB 1473 ACTGTATATATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1532
OY 2137 TAAGCCCGAGCTGTGTATATTGAGGTGATATTATTACGTCTCTGGTCCAGTCTTTTCTG 2196
DB 1533 TCACTGAAATGTTGTAGATCGAGGTGATATTATTAGTCTCTGGTCCAGTCTTCTCTG 1592
OY 2197 GCAATTAACAGTAAGATGTTTAGCAGGTCACTGTTGGGTGAGAGAGTGCATGATC 2256
DB 1593 GCAATTAAGTGAATAATGGCTTAGCAGGTGACCAAGTTGATCCAAAGGAGCTGGTGAGT 1652
OY 2257 ACCAAGCAGGAAAGGAGGAGGAAATAGAGAAATGTGTTGGGTTAGTGAATAATGGCAG 2316
DB 1653 GCCAGGTGGGAGGAGGAGTAGGTAGATAAGGCAATTTGTGTTAAGTTATGAATAATGGC 1712
OY 2317 TGGTGGCCGG 2326
DB 1713 AGGTATCAG 1722
```

RESULT 9

AL120048

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AL120048 610 bp mRNA linear EST 04-SEP-2003
DKEP2761K142_r1 761 (synonym: hamy2) Homo sapiens cDNA clone
DKEP2761K142 5', mRNA sequence.

AL120048

AL120048.1 GI:5925947

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 610)

Ottewaelde, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and

Wiemann, S.

EST (Ottewaelde, et al.)

JOURNAL
COMMENT

Unpublished (1999)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by MediGenomix (Martinsried/Germany) within the cDNA

sequencing consortium of the German Genome Project. No s1 sequence

available.

This clone (DKFZp761K142) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..610

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp761K142"

/tissue type="amygdala"

/dev stage="adult"

/lab_host="DH10B"

/clone_lib="761 (synonym: hamy2)"

/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 19.4%; Score 594.8; DB 1; Length 610;

Best Local Similarity 99.5%; Pred. No. 2e-86;

Matches 607; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1719 AGGAGCACACAGAGGGTAAAGAGGGCTCTCGAAGGACAAAGTTCAGCAGCGGAGAT 1778

Db 1 AGGAGCACACAGAGGGTAAAGAGGGCTCTCGAAGGACAAAGTTCAGCAGCGGAGAT 60

Qy 1779 GAACAGCAGAAAGACCAACAGCAGGAGGACCAAGAACCCAGCCAGTCACAGAGCAGGC 1838

Db 61 GAACAGCAGAAAGACCAACAGCAGGAGGACCAAGAACCCAGCCAGTCACAGAGCAGGC 120

Qy 1839 CACGGTGACACGAACTCAGTGCAGATGGGACAAAGTTCACAAAGAGACCTGAGAAGCG 1898

Db 121 CACGGTGACACGAACTCAGTGCAGATGGGACAAAGTTCACAAAGAGACCTGAGAAGCG 180

Qy 1899 GCAGCAGTCCCTTGGGGCTCTTTAAAGGCTGGGACCAAGCGGATGTTGGATGCTCA 1958

Db 181 GCAGCAGTCCCTTGGGGCTCTTTAAAGGCTGGGACCAAGCGGATGTTGGATGCTCA 240

Qy 1959 AGTGCAACACAGACCCAGTATCATCGGACCAAGTGGGCAACCCCAAGTAAACAATCAGCA 2018

Db 241 AGTGCAACACAGACCCAGTATCATCGGACCAAGTGGGCAACCCCAAGTAAACAATCAGCA 300

Qy 2019 CGGTTCCACACAGGTTCTCTGCGCACCAAGATGTTCTCTTACTCTCATCTCTCCGCCCA 2078

Db 301 CGGTTCCACACAGGTTCTCTGCGCACCAAGATGTTCTCTTACTCTCATCTCTCCGCCCA 360

Qy 2079 AACAGCTCCATGATATATTTCTTGATGCCAGCAAAATGAAATTTGCTTAGAAATTA 2138

Db 361 AACAGCTCCATGATATATTTCTTGATGCCAGCAAAATGAAATTTGCTTAGAAATTA 420

Qy 2139 AGCCGAGCTGTTGATATGAGGTGATATTTACGTCTCTGTCAGTCTTTCTGCGC 2198

Db 421 AGCCGAGCTGTTGATATGAGGTGATATTTACGTCTCTGTCAGTCTTTCTGCGC 480

Qy 2199 AAATAACAGTAAAGATGTTTAGCAGGTCCACTAGTGGGTGAGAGAGTGCATGATCAC 2258

Db 481 AAATAACAGTAAAGATGTTTAGCAGGTCCACTAGTGGGTGAGAGAGTGCATGATCAC 540

Qy 2259 CAAGCAGGAAAGGGAGGAATAG -AGGAATGTGTTTCGGGTTTAAGTGAATGAAATGGCAGT 2317

Db 541 CAAGCAGGAAAGGGAGGAATAGAGGAATGTGTCGGGTTTAAGTGAATGAAATGGCAGT 600

Qy 2318 GGTGCGCGG 2327

Db 601 GGTGCGCGG 610

RESULT 10

LOCUS

DEFINITION

CO724689 820 bp mRNA linear EST 27-JUL-2004
ILLUMIGEN MQO 8605 Katze_MBR Macaca mulatta cDNA clone IBIUW:24007
5' similar to Bases 70 to 720 highly similar to human BCAS1
(Hs.400556), mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Macaca mulatta (rhesus monkey)

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopithecinæ; Macaca.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 820)
Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
Large-scale Rhesus Macaque cDNA Sequencing
Unpublished (2003)
Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2004.01.16. 530 Q20 bases. Contains sequence similar
to Alu Library Preparation: Prof. Michael Katze Lab at University
of Washington DNA Sequencing: Illumigen Biosciences Inc. For
further information, see http://www.macaque.org

PCR Primers

FORWARD: CCCTCACTAAAGGGAACAAA

BACKWARD: CACTATAGGCGAATTGGGTA

Insert Length: 820 Std Error: 0.00

Plate: CL000089 row: G column: 09

Seq primer: CCCTCACTAAAGGGAACAAA

POLYA=No.

Location/Qualifiers

1..820

/organism="Macaca mulatta"

/mol_type="mRNA"

/strain="Indian"

/db_xref="taxon:9544"

/clone="IBIUW:24007"

/sex="female"

/dev stage="adult"

/lab_host="E. coli SOLR"

/clone_lib="Katze_MBR"

/note="Organ: brain; Vector: Uni-ZAP XR; Site 1: EcoR I;
Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis
kit (catalog #200400) and ZAP-CDNA Gigapack III Gold
Cloning Kit (Catalog #200450)"

Query Match 19.3%; Score 592.6; DB 7; Length 820;

Best Local Similarity 93.3%; Pred. No. 4.3e-86;

Matches 652; Conservative 0; Mismatches 44; Indels 3; Gaps 3;

Qy 1667 AACATCATGATCCACAGAAAAGACTATCACACCGCAGAGCCTGAACCAACAGGAGCAC 1726

Db 68 ACATCATCATGATCCGCTCAAAAAGACTATCCACACCCCTTACCTAGAGAAAACAGGAGCGC 127

Qy 1727 CACAGAGGGTAAAGAGGGCTCTCGAAGGACAAAGAGTTCAGCAGCCGAGATGAACAAGC 1786

Db 128 CCCAGAGGGCAAGAGGGCTCTCGAAGGACAAAGAGTTCAGCAGTCAGATGAACAAAC 187

Qy 1787 AGAAGAGCAACAGCAGGAAGCCAAAGAACCCAGCCAGTGCACAGAGCAGGCCAGGTGG 1846

Db 188 AGAAGAGCAACAGCAGGAAGCCAAAGAACCCAGCCAGTGCACAGAGCAGGCCAGGTGG 247

Qy 1847 ACACGAATCTACTCGAGNATGGGACAGCTCCAAAAGAGACCTCAGAGGCGGACGAGT 1906

Db 248 ACACGAATCTACTCGAGNATGGGACAGCTCCAAAAGAGACCTCAGAGGCGGACGAGT 307

QY 1907 CCCTTGGGGGCTTCTTTAAAGCCCTGGGACCAAGCGGATGTTGATGCTCAAGTGCAA 1966
| | | | |
Db 308 CCCTTGGGGGCTTCTTTAAAGCCCTGGGACCAAGCGGATGTTGATGCTCAAGTGCAA 367
| | | | |
QY 1967 CAGACCCAGTATCCATCGGACGAGTTGGCAACCCCAAGTAACAATCAGCACGGTTCCC 2026
| | | | |
Db 368 CAGACCCAGTATCCATCGGACGAGTTGGCAATCCCAAGTAACAATCAGCACGGTTCCC 427
| | | | |
QY 2027 ACCAGGTTCTCTGCGCCACAAGATGTTCTCTTACTCCATCTCTCTCCCAACAGCT 2086
| | | | |
Db 428 A-CAGGTTCTCTGCGCCACAAGATGTTCTCTTACTCCATCTCTCTCTCAACATGTT 486
| | | | |
QY 2087 CCATGTATATTTCTCTGATGGCCAGCAATGAATTTCTGCTAGAAATTAAGCCCGAG 2146
| | | | |
Db 487 CCATGTATATTTCTCTGATAGCCAGCAATGAATTTCTGCTAGAAATTAAGCCCGAG 546
| | | | |
QY 2147 CTGTGTATATTAGGCTGATTATTTAGTCTCTCTGCTCAGTCTTTCTGGCAATAACA 2206
| | | | |
Db 547 CTGTGTATATTAGGCTGATTATTTAGTCTCTCTGCTCAGTCTTTCTGGCAATCACT 606
| | | | |
QY 2207 GT-AAAGATGTTTACAGGTCACCTAGTTGGGTCAGAGAGTCGATGATCACCAAGCAG 2265
| | | | |
Db 607 GTAAAGATGTTTACAGGTCACCTACTTGGTGGAGAGTTGATGATCACCAAGTGG 666
| | | | |
QY 2266 GAAAGGAGGATAGAGAAATGTTTCGGGTTAAGTATGATG-AAAATGGCAGTGTGGCC 2324
| | | | |
Db 667 GAAAGGAGGATAGAGAAATGTTTCGGGTTAAGTATGATGAAATAATGGAAGTGTGGCC 726
| | | | |
QY 2325 GGGCTGTGCTCTCGCTCTGCTGTAATCTCAGCACATTTGGG 2363
| | | | |
Db 727 GGGTGTGCTCTCGCTCTGCTGTAATCTCAGCACATTTTGG 765
| | | | |

RESULT 11
BG702473
LOCUS 602684537F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4817000 5',
DEFINITION mRNA sequence.
ACCESSION BG702473
VERSION BG702473.1 GI:13973847
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
DNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10716 row: j column: 09
High quality sequence stop: 658.
Location/Qualifiers
1. 1076
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4817000"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_95"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI

(gtcag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTN-3', size-selected for average
insert size 2.5 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this
is a NIH_MGC Library."

ORIGIN

Query Match 19.1%; Score 586.4; DB 4; Length 1076;
Best Local Similarity 98.0%; Pred. No. 4.1e-85;
Matches 546; Conservative 0; Mismatches 7; Indels 6; Gaps 5;
QY 733 GACAAGGTCCTCTCTGCGCCAGGATCCAC-GCTTCTCCACCTGAGACAGGGGAGC 791
| | | | |
Db 6 GCGAAGGTCCTCTCTGCGCCAGGATCCACGGCTTCTCCACCTGAGACAGGGGAGC 65
| | | | |
QY 792 AGGAGGAGAGCTCCCTCCAGCCCAAGGATCCAGCTTTTGGCAATTTCTTCAAGCT 851
| | | | |
Db 66 AGGAGGAGAGCTCCCTCCAGCCCAAGGATCCAGCTTTTGGCAATTTCTTCAAGCT 125
| | | | |
QY 852 GGACAAGGACAGGAAAGGTGCCAGGTGACAGCCCAAGAGCAAGAGGCGAGCA 911
| | | | |
Db 126 GGACAAGGACAGGAAAGGTGCCAGGTGACAGCCCAAGAGCAAGAGGCGAGCA 185
| | | | |
QY 912 TCAAGACAAGGTGATGAGTTCTTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCAGG 971
| | | | |
Db 186 TCAAGACAAGGTGATGAGTTCTTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCAGG 245
| | | | |
QY 972 GAAGGACATAGTTGACGGCAAGGAAAAA-GAAGGACAAGAACTTGGAACTCGGATGCT 1030
| | | | |
Db 246 GAAGGACATAGTTGACGGCAAGGAAAAA-CGAAGGACAAGAACTTGGAACTCGGATGCT 305
| | | | |
QY 1031 CTGTCTCTGGGGACCCAGAGGACTGGAGACTGCAAGAGAGATTCGCCAGGACAGCTA 1090
| | | | |
Db 306 CTGTCTCTGGGGACCCAGAGGACTGGAGACTGCAAGAGAGATTCGCCAGGACAGCTA 365
| | | | |
QY 1091 TAGCAGAGAAATAAATAATTCATCATGAGTTCTTTAAAACTCTGGTTTCACTAACAAAG 1150
| | | | |
Db 366 TAGCAGAGAAATAAATAATTCATCATGAGTTCTTTAAAACTCTGGTTTCACTAACAAAG 425
| | | | |
QY 1151 CTGAAACAAAAAGGACCCAGAGACACCGGTGCTGAAAAGTCAACCACTTCAGCTG 1210
| | | | |
Db 426 CTGAAACAAAAAGGACCCAGAGACACCGGTGCTGAAAAGTCAACCACTTCAGCTG 485
| | | | |
QY 1211 ACCTTTAAGTCAGACAAAGCCAACTTTATCATCCAGGACACCAAGGGCTGCAAGAAAT 1270
| | | | |
Db 486 ACCTTTAAGTCAGACAAAGCCAACTTTATCATCCAGGACACCAAGGGCTGCAAGAAAT 545
| | | | |
QY 1271 CCAAAGGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTGAACTGCGAAGG 1330
| | | | |
Db 546 CCAAAGGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCC--TGAACCTGCGAAGG 603
| | | | |
QY 1331 AAGGACCAAGAGAAATCAGACCCACTCTTCGCTCT--GGGCAAACTGTTTGGAA 1388
| | | | |
Db 604 AAGG-ACCAAGGAGAAATCAGACCCCACTCTCTGCTCTGCGGCAAACTGTTTGGAA 661
| | | | |

RESULT 12

BP199106
LOCUS BP199106 Sugano cDNA library, brain Homo sapiens EST 14-SEP-2004
DEFINITION ADR06635, mRNA sequence.
ACCESSION BP199106
VERSION BP199106.1 GI:52046726
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yeuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..598
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="ADR06635"
/tissue type="brain"
/clone_lib="Sugano cDNA library, brain"

ORIGIN
Query Match 19.0%; Score 583.4; DB 5; Length 598;
Best Local Similarity 99.3%; Pred. No. 1.4e-84;
Matches 595; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 670 CTTCCGGTGGCAGCTGGACCGGGGAGGACACAGATAAAACCCAGGGGACGCCCCGGCC 729
Db 1 CTTCCGGTGGCAGCTGGACCGGGGAGGACACAGATAAAACCCCA-GGCACGCCCCGGCC 59
Qy 730 CAAGACAGGTCCTCTCTGCCCGCAGGATCCCGCTTCTCCACCTGACACAGGGGGA 789
Db 60 CAAGACAGGTCCTCTCTGCCCGCAGGATCCCGCTTCTCCACCTGACACAGGGGGA 119
Qy 790 GCAGGAGAGAGCTCCCTCCAAAGCCCAAGACTCCAGCTTTTGTGACAAATCTTCAAG 849
Db 120 GCAGGAGAGAGCTCCCTCCAAAGCCCAAGACTCCAGCTTTTGTGACAAATCTTCAAG 179
Qy 850 CTGCAAGAGGACAGGAAAGGTGCCAGTGACAGCCAAAGGAGCCAAAGAGGGCAGAG 909
Db 180 CTGCAAGAGGACAGGAAAGGTGCCAGTGACAGCCAAAGGAGCCAAAGAGGGCAGAG 239
Qy 910 CATCAAGACAGGTGGATGAGGTCTCGCTTATCAGGGCAGTCCGATGATGTCCTTGCA 969
Db 240 CATCAAGACAGGTGGATGAGGTCTCGCTTATCAGGGCAGTCCGATGATGTCCTTGCA 299
Qy 970 GGGAGGACATAGTTGACGGCAAGGAAAGGACAGCAAGAACTTGGAACTGCGGATTGC 1029
Db 300 GGGAGGACATAGTTGACGGCAAGGAAAGGACAGCAAGAACTTGGAACTGCGGATTGC 359
Qy 1030 TCTGTCCCTGGGACCCAGAGGACTGGAGACTGCAAAAGGACGATTCAGGCGCAGCAGCT 1089
Db 360 TCTGTCCCTGGGACCCAGAGGACTGGAGACTGCAAAAGGACGATTCAGGCGCAGCAGCT 419
Qy 1090 ATAGCAGAGATAATAATTCATCATGATTTCTTTAAACTCTGTTTCACTTAACAAA 1149
Db 420 ATAGCAGAGATAATAATTCATCATGATTTCTTTAAACTCTGTTTCACTTAACAAA 479
Qy 1150 GCTGAAACAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACACTTCAGCT 1209
Db 480 GCTGAAACAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACACTTCAGCT 539
Qy 1210 GACCTTAAGTCAGACAAAGCCAACTTTTATCCCAAGGAGACCAAGGGGCTGGCAAGAA 1268
Db 540 GACCTTAAGTCAGACAAAGCCAACTTTTATCCCAAGGAGACCAAGGNGCTGGNCAGAA 598

RESULT 13
BP342407
LOCUS BP342407 Sugano cDNA library, stomach mucosa Homo sapiens cDNA
DEFINITION clone STM04831, mRNA sequence.
ACCESSION BP342407
VERSION BP342407.1 GI:52272209
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 583)
AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yeuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..583
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="STM04831"
/tissue type="stomach mucosa"
/clone_lib="Sugano cDNA library, stomach mucosa"

ORIGIN
Query Match 19.0%; Score 581.4; DB 5; Length 583;
Best Local Similarity 99.8%; Pred. No. 3e-84;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 146 AGCCAGGACCTGGGACGACGCACTGGAGACCCAGGACCCCTGTGCAGGAGCAGCTCCGGG 205
Db 1 AGCCAGGACCTGGGACGACGCACTGGAGACCCAGGACCCCTGTGCAGGAGCAGCTCCGGG 60
Qy 206 TGACACAGGAGGACTGAAGATACCTCCACAGGGGCTCAGCAGGAGCAATGGTAAACAAA 265
Db 61 TGACACAGGAGGACTGAAGATACCTCCACAGGGGCTCAGCAGGAGCAATGGTAAACAAA 120
Qy 266 TGAGTGTTCGCCAAAGAGTTGAAGACCAAGAGAAATGAACAGAGAGAGACTTACCAAG 325
Db 121 TGAGTGTTCGCCAAAGAGTTGAAGACCAAGAGAAATGAACAGAGAGAGACTTACCAAG 180
Qy 326 ACACAGGCTCTGCTCTGACCGGGTTCAGTGGTGGTTCGACCCACACAGCTTCAGCACT 385
Db 181 ACACAGGCTCTGCTCTGACCGGGTTCAGTGGTGGTTCGACCCACACAGCTTCAGCACT 240
Qy 386 TAGAGGAAGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGCG 445
Db 241 TAGAGGAAGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGCG 300
Qy 446 AGACACAGGAGTAAGTGTGTCGGATGCCAAGGAAAGAAATCTTGGGAAAGAGGCCA 505
Db 301 AGACACAGGAGTAAGTGTGTCGGATGCCAAGGAAAGAAATCTTGGGAAAGAGGCCA 360
Qy 506 AACCCGGGACACAGCTGCTAAATCTGTTTTTCTTGATGCTCTCTCGGCTGTACCAG 565
Db 361 AACCCGGGACACAGCTGCTAAATCTGTTTTTCTTGATGCTCTCTCGGCTGTACCAG 420
Qy 566 GACGTACCGGAGACCAAGCCGACAGATTCATCCCTTGGATCAGTGAAGCTTGTACGCT 625
Db 421 GACGTACCGGAGACCAAGCCGACAGATTCATCCCTTGGATCAGTGAAGCTTGTACGCT 480
Qy 626 CCAATAAAGTCTCAGCGCAACAAAGACCAAGTGAAGCTGGACACTTCCGGTGGCAGCTG 685
Db 481 CCAATAAAGTCTCAGCGCAACAAAGACCAAGTGAAGCTGGACACTTCCGGTGGCAGCTG 540
Qy 686 GACCCGGGACGACACAGATAAAACCCAGGACGCCCCCGGC 728
Db 541 GACCCGGGACGACACAGATAAAACCCAGGACGCCCCCGGC 583

RESULT 14
BM702087
LOCUS BM702087
DEFINITION UI-E-CQ1-aey-i-21-0-UI.r1 UI-E-CQ1 Homo sapiens cDNA clone

UI-E-COI-aey-i-21-0-UI 5', mRNA sequence.
BM702087
BM702087.1 GI:19015345
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 589)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PubMed 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA sequence: 390-589, >ALU
Seq primer: M13 Reverse
Location/Qualifiers
1..589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-COI-aey-i-21-0-UI"
/tissue_type="optic nerve"
/dev_stages="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-COI"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-COI is a normalized cDNA library containing the following tissue(s): optic nerve. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCATTAAGTG. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."
ORIGIN
Query Match 18.9%; Score 581; DB 4; Length 589;
Best Local Similarity 99.2%; Pred. No. 3.5e-84;
Matches 584; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1932 GGGACCAAGCGGATGTTGGATGCTCAAGTGCACACAGACCCAGATCATCCGACCAAGT 1991
DB 1 GGGACCAAGCGGATGTTGGATGCTCAAGTGCACACAGACCCAGATCATCCGACCAAGT 60
QY 1992 TGGCAAAACCCAGTAAGTAACAAATCAGACGGTTCCACCAAGGTTCTTCGCCACCAAGATG 2051
DB 61 TGGCAAAATCCAGTAAGTAACAAATCAGACGGTTCCACCAAGGTTCTTCGCCACCAAGATG 120
QY 2052 TGTTCCTCTTACTCCTCTCTCCCAACAGCGTCCATGTATATATTCTTCTGTATGCC 2111

Db 121 TGTTCCTCTTACTCCTCTCTCCCAACAGCGTCCATGTATATATTCTTCTGTATGCC 180
QY 2112 AGCAAAATGAAATTTGCCTAGAAATTAAGCCGAGCTGTTGATATTGAGGTGTTATT 2171
Db 181 AGCAAAATGAAATTTGCCTAGAAATTAAGCCGAGCTGTTGATATTGAGGTGTTATT 240
QY 2172 TACGTCCTCTGTCAGTCTCTTTCTGGCAAAATTAACAGTAAGATGTTTAGCAGGTCCACT 2231
Db 241 TACGTCCTCTGTCAGTCTCTTTCTGGCAAAATTAACAGTAAGATGTTTAGCAGGTCCACT 300
QY 2232 AGTTGGGTTCAGAAAGATGCGATGATCACCAGCAGGAAAGGAGGGAATAGAGGAATGTGT 2291
Db 301 AGTTGGGTTCAGAAAGATGCGATGATCACCAGCAGGAAAGGAGGGAATAGAGGAATGTGT 360
QY 2292 TCGGTTAAAGTGAATAAATGCGAGTGTGCGCGGCGTGTGGTCTCGCCTGTAATCT 2351
Db 361 TCGGTTAAAGTGAATAAATGCGAGTGTGCGCGGCGTGTGGTCTCGCCTGTAATCT 420
QY 2352 CAGCACATTTGGGAGCGGCGAGGCGAGTGTGATCACCTGAGTTCAGGAGTTCAGACTAGCT 2411
Db 421 CAGCACATTTGGGAGCGGCGAGGCGAGTGTGATCACCTGAGTTCAGGAGTTCAGACTAGCT 480
QY 2412 GGCCAAACATCATGAAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGGATGGTGC 2471
Db 481 GGCCAAACATCATGAAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGGATGGTGC 540
QY 2472 ACACACTGTAGTCCAGTCTACTCGGAGCGCAACGACGAGAACCGCT 2520
Db 541 ACACACTGTAGTCCAGTCTACTCGGAGTCCCAACGACGAGAACCGCT 589
RESULT 15
BX486692
LOCUS
DEFINITION DKFZp686P08252_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION BX486692
VERSION BX486692.1 GI:31950607
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (Bases 1 to 583)
AUTHORS Bahr,A., Lauber,J., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
EST (Bahr,A., Lauber,J., Mewes,H.W., Weil,B., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.
No s1 sequence available.
This clone (DKFZp686P08252) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686P08252"
/dev_stages="adult"
/lab_host="DH10B"
/note="Vector: pTriplex2; (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: Sf1A; Site_2: Sf1B; cDNA-collection"

| | | | | | |
|-----------------------|-----|---|--------------------|-----------|-------------|
| Query Match | | 18.7%; | Score 573.2; | DB 5; | Length 583; |
| Best Local Similarity | | 99.5%; | Pred. NO. 6.5e-83; | | |
| Matches 575; | | Conservative 0; | Mismatches 3; | Indels 0; | Gaps 0; |
| Qy | 127 | GGCAGGAGCTCCACCGACAGCCAGGCACTGGGAGCAGCGCACTGGAGACCCAGGACCCCT | 186 | | |
| Db | 6 | GGAGCCAGGCTCCACCGACAGCCAGGCACTGGGAGCAGCGCACTGGAGACCCAGGACCCCT | 65 | | |
| Qy | 187 | GTGAGGAGAGCAGCTCCGCGGTGACACGAGGGACTGAAGATACTCCACAGGGGTCTCAGCA | 246 | | |
| Db | 66 | GTGAGGAGAGCAGCTCCGCGGTGACACGAGGGACTGAAGATACTCCACAGGGGTCTCAGCA | 125 | | |
| Qy | 247 | GGAGCAATGGGTAAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGAATCAACCA | 306 | | |
| Db | 126 | GGAGCAATGGGTAAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGAATGAACCA | 185 | | |
| Qy | 307 | GAAGCAGAGACTTACCAGGACAAACGCTCTCTCTGAAACGGGGTTCAGTGGTGGTGTCTG | 366 | | |
| Db | 186 | GAAGCAGAGACTTACCAGGACAAACGCTCTCTCTGAAACGGGGTTCAGTGGTGGTGTCTG | 245 | | |
| Qy | 367 | ACCCACACAGTTTACAGCACTTAGAGGAAGTTCGACTTGGGAATAAGTGTCAAGACGGATAAT | 426 | | |
| Db | 246 | ACCCACACAGTTTACAGCACTTAGAGGAAGTTCGACTTGGGAATAAGTGTCAAGACGGATAAT | 305 | | |
| Qy | 427 | GTGGCCACTTCTTCCCGCAGACAAACGAGATAAGTGTGTGGGATGCCAACGGAAG | 486 | | |
| Db | 306 | GTGGCCACTTCTTCCCGCAGACAAACGAGATAAGTGTGTGGGATGCCAACGGAAG | 365 | | |
| Qy | 487 | AATCTTGGGAAAGAGCCAAACCCGAGGACCAAGCTGCTAAATCTCGTTTTTTCTTGATG | 546 | | |
| Db | 366 | AATCTTGGGAAAGAGGCCAAACCCGAGGACCAAGCTGCTAAATCTCGTTTTTTCTTGATG | 425 | | |
| Qy | 547 | CTCTCTCGGCTGTACACAGGACGTACCGGAGACCAAGCGCAGATTATCCCTTGGATCA | 606 | | |
| Db | 426 | CTCTCTCGGCTGTACACAGGACGTACCGGAGACCAAGCGCAGATTATCCCTTGGATCA | 485 | | |
| Qy | 607 | GTGAAGCTTGATGTTCAGCTCCAATAAAGCTCCAGCGAAACAAAGACCCCAAGTGAGAGCTGG | 666 | | |
| Db | 486 | GTGAAGCTTGATGTTCAGCTCCAATAAAGCTCCAGCGAAACAAAGACCCCAAGTGAGAGCTGG | 545 | | |
| Qy | 667 | ACACTTCGGTGGCAGCTGGACCGGGGAGGACACAGA | 704 | | |
| Db | 546 | ACACTTCGGTGGCAGCTGGACCGGGGAGGACACAGA | 583 | | |

Search completed: July 3, 2005, 16:40:15
Job time : 9707.27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:48:07 ; Search time 5109.25 Seconds
(without alignments)
11428.001 Million cell updates/sec

Title: US-08-731-499-7
Perfect score: 1205
Sequence: 1 GCGCGGTGAGTCCGCCCCC.....TCAGTGTGTTGTTAAGTGG 1205

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:

- 1: gb_ba:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_on:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 1201 | 99.7 | 1205 | 6 | AR070330 Sequence |
| 2 | 1201 | 99.7 | 1205 | 6 | BD085731 Genes fro |
| c 3 | 844.6 | 70.1 | 122764 | 9 | AL109984 Human DNA |
| c 4 | 742.8 | 61.6 | 173718 | 9 | AC073575 Homo sapi |
| 5 | 739.8 | 61.4 | 892 | 9 | X94910 Homo sapien |
| 6 | 738.6 | 61.3 | 889 | 6 | C0728979 Sequence |
| 7 | 635.6 | 52.7 | 786 | 9 | CR541667 |
| 8 | 545 | 45.2 | 1258 | 10 | BC017125 |
| 9 | 496.6 | 41.2 | 1139 | 6 | AX827789 |
| 10 | 496.6 | 41.2 | 1139 | 10 | RNU36482 |
| 11 | 479.4 | 39.8 | 1084 | 10 | RNERP29PR |
| 12 | 404.8 | 33.6 | 438 | 6 | C0692426 |
| 13 | 403 | 33.4 | 464 | 11 | G25553 |
| 14 | 403 | 33.4 | 464 | 11 | G27410 |
| 15 | 402.2 | 33.4 | 500 | 6 | AX887897 |
| 16 | 402.2 | 33.4 | 500 | 6 | BD027507 Sequence |
| c 17 | 358.8 | 29.8 | 140936 | 10 | AC129215 |
| 18 | 295.8 | 24.5 | 9664 | 10 | AY004254 |
| c 19 | 295.8 | 24.5 | 216476 | 2 | AC095682 Rattus no |

| | | | | | | |
|------|-------|------|--------|---|----------|-----------|
| 20 | 279.2 | 23.2 | 450 | 6 | C0051538 | Sequence |
| 21 | 279.2 | 23.2 | 450 | 6 | C0065889 | Sequence |
| 22 | 279.2 | 23.2 | 450 | 6 | C0093637 | Sequence |
| 23 | 279.2 | 23.2 | 450 | 6 | CQ132410 | Sequence |
| 24 | 279.2 | 23.2 | 450 | 6 | CQ170979 | Sequence |
| 25 | 279.2 | 23.2 | 450 | 6 | CQ200125 | Sequence |
| 26 | 279.2 | 23.2 | 450 | 6 | CQ215610 | Sequence |
| 27 | 279.2 | 23.2 | 450 | 6 | CQ254208 | Sequence |
| 28 | 279.2 | 23.2 | 450 | 6 | CQ291221 | Sequence |
| 29 | 279.2 | 23.2 | 450 | 6 | CQ328331 | Sequence |
| 30 | 276 | 22.9 | 393 | 6 | CQ732783 | Sequence |
| 31 | 252.4 | 20.9 | 255 | 6 | CQ686591 | Sequence |
| 32 | 252 | 20.9 | 255 | 6 | CQ678911 | Sequence |
| 33 | 238.6 | 19.8 | 711 | 5 | AJ720735 | Gallus ga |
| c 34 | 194.8 | 16.2 | 244208 | 2 | AC095446 | Rattus no |
| 35 | 194.8 | 16.2 | 269849 | 2 | AC131616 | Rattus no |
| 36 | 191.2 | 15.9 | 363 | 6 | AX892436 | Sequence |
| 37 | 191.2 | 15.9 | 363 | 6 | BD027969 | Sequence |
| 38 | 186.8 | 15.5 | 197 | 6 | CQ674978 | Sequence |
| 39 | 140 | 11.6 | 210 | 6 | CQ659512 | Sequence |
| 40 | 99.6 | 8.3 | 119 | 6 | CQ056512 | Sequence |
| 41 | 99.6 | 8.3 | 119 | 6 | CQ075782 | Sequence |
| 42 | 99.6 | 8.3 | 119 | 6 | CQ106764 | Sequence |
| 43 | 99.6 | 8.3 | 119 | 6 | CQ145417 | Sequence |
| 44 | 99.6 | 8.3 | 119 | 6 | CQ180859 | Sequence |
| 45 | 99.6 | 8.3 | 119 | 6 | CQ205231 | Sequence |

ALIGNMENTS

RESULT 1
AR070330
LOCUS AR070330 1205 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 7 from patent US 5892010.
ACCESSION AR070330
VERSION AR070330.1 GI:7221218
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1205)
AUTHORS Gray,J., Collins,C., Hwang,S.-i., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20Q13 amplicon and their uses
JOURNAL Patent: US 5892010-A 7 06-APR-1999;
FEATURES Location/Qualifiers
source 1..1205
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 99.7%; Score 1201; DB 6; Length 1205;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGCGGTGAGTCCGCCCCCAGTCACTGACCGCTGACTCGGGGGCTTCTCCACTATC 60
Db 1 GCGCGGTGAGTCCGCCCCCAGTCACTGACCGCTGACTCGGGGGCTTCTCCACTATC 60
Qy 61 GCTTACCTACCTCCCTCTGAGGAAACCGGGGATATGCTCGCTGTCGCCCGCCGCGC 120
Db 61 GCTTACCTACCTCCCTCTGAGGAAACCGGGGATATGCTCGCTGTCGCCCGCCGCGC 120
Qy 121 ATTTCTCTCCCGGTGCTTCCCTTCTCTGGGCTTCTGCTCCTCTCGCTCCGATGCG 180
Db 121 ATTTCTCTCCCGGTGCTTCCCTTCTCTGGGCTTCTGCTCCTCTCGCTCCGATGCG 180
Qy 181 GGCAAGCGCTTGACACCAAGCGCCCTTCCCTGGATACGGTCACTTTCTACAAGGTCA 240
Db 181 GGCAAGCGCTTGACACCAAGCGCCCTTCCCTGGATACGGTCACTTTCTACAAGGTCA 240
Qy 241 TTCCCAAAAGCAAGTGTGCTGTGGTGAAGTTTCGACACCCAGTACCCCTACGGTGAGAAGCA 300

misc_feature /note="AluSx repeat: matches 1. .302 of consensus"
complement(2259. .2786)
/note="match: GSS: Em:AQ549766"
2391. .2410
repeat_region /note="4.0 copies 5 mer CAAAA 40% conserved"
2417. .2432
repeat_region /note="5.3 copies 3 mer AAT 32% conserved"
2455. .2464
repeat_region /note="10.0 copies 1 mer T 20% conserved"
2482. .2491
repeat_region /note="3.3 copies 3 mer TTA 20% conserved"
2542. .2561
repeat_region /note="2.2 copies 9 mer TTGTAGGC 31% conserved"
2544. .2657
repeat_region /note="MER69 repeat: matches 17. .147 of consensus"
complement(2592. .2794)
misc_feature /note="match: GSS: Em:AQ467488"
2685. .2731
repeat_region /note="MER69 repeat: matches 2476. .2523 of consensus"
2803. .3228
misc_feature /note="match: GSS: Em:AQ332266"
2859. .2943
repeat_region /note="MIR repeat: matches 61. .147 of consensus"
3001. .3014
repeat_region /note="2.0 copies 7 mer AAAATGG 28% conserved"
3389. .3510
repeat_region /note="MIR repeat: matches 2. .119 of consensus"
3398. .3409
repeat_region /note="2.0 copies 6 mer GTAATG 24% conserved"
3511. .3791
repeat_region /note="AluY repeat: matches 1. .311 of consensus"
3621. .3647
repeat_region /note="1.9 copies 14 mer CCTGTAGTCCCAGG 54% conserved"
3762. .3791
repeat_region /note="30.0 copies 1 mer A 60% conserved"
3792. .3917
repeat_region /note="MIR repeat: matches 119. .258 of consensus"
complement(3955. .4264)
repeat_region /note="AluSx repeat: matches 1. .309 of consensus"
3955. .3981
repeat_region /note="27.0 copies 1 mer T 27% conserved"
3991. .4006
repeat_region /note="2.0 copies 8 mer TATTGCTC 32% conserved"
4206. .4215
repeat_region /note="2.5 copies 4 mer CTGC 20% conserved"
4275. .4323
repeat_region /note="9.8 copies 5 mer TTCT 54% conserved"
4279. .4324
repeat_region /note="46.0 copies 1 mer T 47% conserved"
complement(4295. .4605)
repeat_region /note="AluY repeat: matches 1. .311 of consensus"
4680. .4696
repeat_region /note="4.2 copies 4 mer ATGT 25% conserved"
complement(4719. .5094)
repeat_region /note="L1M4 repeat: matches 3083. .3493 of consensus"
4723. .4733
repeat_region /note="2.2 copies 5 mer TAATT 22% conserved"
4879. .4890
repeat_region /note="2.0 copies 6 mer TCTGCA 24% conserved"
5082. .5094
repeat_region /note="13.0 copies 1 mer T 26% conserved"
complement(5159. .5450)
repeat_region /note="AluSg repeat: matches 10. .300 of consensus"
5159. .5178
repeat_region /note="20.0 copies 1 mer T 31% conserved"
5242. .5253
repeat_region /note="2.0 copies 6 mer ACCTCC 24% conserved"
5505. .5514
repeat_region /note="3.3 copies 3 mer TTA 20% conserved"
5514. .5548
repeat_region /note="MIR repeat: matches 110. .144 of consensus"
complement(5711. .5887)
repeat_region /note="MER45A repeat: matches 1. .178 of consensus"

repeat_region 5982. .5993
/note="2.0 copies 6 mer CATTCT 24% conserved"
6019. .6511
misc_feature /note="match: GSS: Em:AQ473868"
6490. .6500
repeat_region /note="3.7 copies 3 mer TGC 22% conserved"
complement(6634. .6722)
repeat_region /note="MIR repeat: matches 11. .104 of consensus"
6740. .7042
repeat_region /note="AluJb repeat: matches 1. .302 of consensus"
7023. .7042
repeat_region /note="20.0 copies 1 mer A 40% conserved"
7164. .7177
repeat_region /note="3.5 copies 4 mer TTAT 28% conserved"
complement(7167. .7480)
repeat_region /note="AluSp repeat: matches 1. .298 of consensus"
7240. .7273
repeat_region /note="2.0 copies 17 mer TGGCAACTTCCTCC 68% conserved"
complement(7488. .7565)
repeat_region /note="MIR repeat: matches 74. .156 of consensus"
7488. .7498
repeat_region /note="3.7 copies 3 mer ATT 22% conserved"
7607. .7616
repeat_region /note="2.5 copies 4 mer AAAC 20% conserved"
complement(7623. .7917)
repeat_region /note="AluY repeat: matches 1. .295 of consensus"
7623. .7636
repeat_region /note="14.0 copies 1 mer T 28% conserved"
8078. .8092
repeat_region /note="3.0 copies 5 mer GAAAA 30% conserved"
complement(8471. .8780)
repeat_region /note="AluSg repeat: matches 1. .307 of consensus"
8482. .8494
repeat_region

Query Match 70.1%; Score 844.6; DB 9; Length 122764;
Best Local Similarity 88.7%; Pred. No. 3.2e-228;
Matches 1058; Conservative 4; Mismatches 110; Indels 21; Gaps 13;

Qy 30 TGACCGCTGACTCGGGCGTTCTCCACTATCGCTTACTACTCTCCCTCTGCAGAACCCG 89
Db 93229 TGACCGCTGACTTGGGGGTTCTCCACTACGCTTCCCACTCCCGTCTCTCCCTCCCTCTCC 93170
Qy 90 GCGATATGGCTGCGCTGTGCCCGCGCCGCAATTCTCTCCCGCTGCTT-CCCTTCTCC 148
Db 93169 GCGATGTGGCTGCGCGCGTGCCTGCTCTCTCTCCCTGCGCTTCCCTCTCC 93110
Qy 149 TGGGCTTCTGCTCTCTCCGCTCGGATCGGCGAGGCGCTGCACACCAAG-CGCGCC 207
Db 93109 TGGGCTTCTGCTCTCTCCGCTCGGATCGGCGAGGCGCTGCACACCAAGCGCGCC 93050
Qy 208 TTCCCTGGATACGGTCACTTCTACAAGTCAATTCCTCCAAAGCAAGTTCGTC-TGGTGA 266
Db 93049 TTCCCTGGATACGGTCACTTCTACAAGTCAATTCCTCCAAAGCAAGTTCGTC-TGGTGA 92990
Qy 267 AGTTTCGACACCCAGTACCCCTACCGGTGAGAGAGAGAGTTCAGGCTCTT-CTGAAA 325
Db 92989 AGTTTCACACCCAGTACCGCTACCGTACGAGAGAGAGTTCAGGCTCTTTCGTGAAA 92930
Qy 326 ACTGGCTTCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTAT-GTGACA 384
Db 92929 ACTGGCTTCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTATGTGTGACA 92870
Qy 385 AGCTGAACATGGAGCTGAGTGAGAAATACAAAGCTGGACAAAGAGAGCTACCCA-TCTTCT 443
Db 92869 AGCTGAACATGGAGCTGAGTGAGAAATACAAAGCTGCTCAAGAGAGCTACCCAGTCTTCT 92810
Qy 444 ACCTCTTCGCGATGGGACTTTGAGAACCCAGTCCCAATACACTGGGGCAGTT-AGGTTG 502
Db 92809 ACCTCTCCGGAGGGGACTTTGAGAACCCAGTCCCAATGATGACTGGGCGAGTTAGGTTG 92750
Qy 503 GAGCATCAGCGCTGGCTGAGGGGCAAGGGGCTACTAGGTAT-----GCCTGG 554
Db 92749 GAGCATCAGCGCTGGCTGAGGGGCAAGGGGCTTAGCTAGCTATGCCAGGTGCGCTGG 92690

standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES

source

Location/Qualifiers

```
1. 173718
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP1-267L14"
complement(6..99)
/rpt_family="AluS"
repeat_region
complement(137..265)
/rpt_family="AluSx"
repeat_region
complement(521..817)
/rpt_family="AluSx"
repeat_region
complement(2270..2533)
/rpt_family="AluSx"
repeat_region
complement(2769..3011)
/rpt_family="AluSx"
repeat_region
complement(3404..3705)
/rpt_family="AluSx"
repeat_region
complement(3909..4120)
/rpt_family="AluSx"
repeat_region
complement(5561..5949)
/rpt_family="AluSx"
repeat_region
complement(6237..6519)
/rpt_family="AluSx"
repeat_region
complement(6530..6814)
/rpt_family="AluSx"
repeat_region
complement(6815..7117)
/rpt_family="AluSx"
repeat_region
complement(7280..7389)
/rpt_family="AluSx"
repeat_region
complement(7390..7703)
/rpt_family="AluSx"
repeat_region
complement(7704..7887)
/rpt_family="AluSx"
repeat_region
complement(8504..8590)
/rpt_family="AluSx"
repeat_region
complement(8881..9172)
/rpt_family="AluSx"
repeat_region
complement(9424..9521)
/rpt_family="AluSx"
repeat_region
complement(10959..11094)
/rpt_family="AluSx"
repeat_region
complement(12307..12408)
/rpt_family="AluSx"
repeat_region
complement(12866..13012)
/rpt_family="AluSx"
repeat_region
complement(14422..14540)
/rpt_family="AluSx"
repeat_region
complement(15362..15432)
/rpt_family="AluSx"
repeat_region
complement(15433..15733)
/rpt_family="AluSx"
repeat_region
complement(15734..15824)
/rpt_family="AluSx"
repeat_region
complement(15867..16158)
/rpt_family="AluSx"
repeat_region
complement(16164..16459)
/rpt_family="AluSx"
repeat_region
complement(16608..16691)
/rpt_family="AluSx"
repeat_region
```

```
STS
19874..20015
/standard_name="SHGC-17261"
STS
20458..20707
/standard_name="WI-19956"
STS
20529..20719
/standard_name="RH80597"
repeat_region
20632..20653
/rpt_family="AT_rich"
repeat_region
25484..25780
/rpt_family="AluY"
repeat_region
26272..26557
/rpt_family="AluSg"
STS
27536..27667
/standard_name="A006C17"
repeat_region
29386..29691
/rpt_family="AluSx"
repeat_region
complement(29836..30125)
/rpt_family="AluSg"
repeat_region
30221..30528
/rpt_family="AluSx"
repeat_region
30535..30825
/rpt_family="AluSx"
repeat_region
complement(31613..31722)
/rpt_family="HY1"
repeat_region
complement(31836..32147)
/rpt_family="AluSx"
STS
32373..32700
/standard_name="SHGC-149403"
```

Query Match 61.6%; Score 742.8; DB 9; Length 173718;

Best Local Similarity 98.1%; Pred. No. 2.8e-199;

Matches 818; Conservative 5; Mismatches 4; Indels 7; Gaps 7;

```
Qy 379 GTGACAGCTGACATGAGCTGAGTGAAGATACAGCTGACAAAGAGAGCTACCCCA- 437
Db 158927 GTGACAGCTGACATGAGCTGAGTGAAGATACAGCTGACAAAGAGAGCTACCCAG 158868
Qy 438 TCTTCTACCTCTCCGGATGGGACCTTGAAGACCCAGTCCCATACACTCGGGCAGTT- 496
Db 158867 TCTTCTACCTCTCCGGATGGGACCTTGAAGACCCAGTCCCATACACTCGGGCAGTTA 158808
Qy 497 AGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTACTAGTATGCTCGG-T 555
Db 158807 AGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTACTAGTATGCTCGGTT 158748
Qy 556 GCCTGCCTGTATACGACGCCCTGGCGGGAGTTTCATCAGGGCCTCTCGTGTGGAGG-CC 614
Db 158747 GCCTGCCTGTATACGACGCCCTGGCGGGAGTTTCATCAGGGCCTCTCGTGTGGAGGCC 158688
Qy 615 GCCAGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAG- 673
Db 158687 GCCAGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAGA 158628
Qy 674 AGTGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-C 732
Db 158627 AGTGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTTCC 158568
Qy 733 CAGCATCAGAGATGACACGGATCGCAGGCTGATTGAGAAGACAGAGTGTGAC-CGC 791
Db 158567 CAGCATCAGAGATGACACGGATCGCAGGCTGATTGAGAAGACAGAGTGTGACCGGA 158508
Qy 792 AGAAGGAGGAGCTCCAGAAGAGCTTTAAACATCTCTGACTGCCTTCAGAAAGAGGGGGCG 851
Db 158507 AGAAGGAGGAGCTCCAGAAGAGCTTTAAACATCTCTGACTGCCTTCAGAAAGAGGGGGCG 158448
Qy 852 AGAAAGAGGAGCTGTAAAGAGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAG 911
Db 158447 AGAAAGAGGAGCTGTAAAGAGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAG 158388
Qy 912 GGGAGAGTTAACCTGCTGCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAG 971
Db 158387 GGGAGAGTTAACCTGCTGCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAG 158328
```


LOCUS CQ728979 889 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 14913 from Patent WO02068579.
ACCESSION CQ728979
VERSION CQ728979.1 GI:42299268
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,B.W.
TITLES Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 14913 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..889
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 61.3%; Score 738.6; DB 6; Length 889;
Best Local Similarity 97.9%; Pred. No. 2.9e-198;
Matches 870; Conservative 3; Mismatches 4; Indels 12; Gaps 12;
Qy 86 CCCGGCATATGGCTCGCGCTGTCGCCCGCGCGCATTTCTCCCGCTGCTT-CCCTT 144
Db 1 CCCGGCATATGGCTCGCGCTGTCGCCCGCGCGCATTTCTCCCGCTGCTTCCCTT 60
Qy 145 CTCCTGGGCTTCCTGCTCCTCTCCGCTCCGATGCGCGCGGCTGCACACCAA-GGC 203
Db 61 CTCCTGGGCTTCCTGCTCCTCTCCGCTCCGATGCGCGCGGCTGCACACCAAGGC 120
Qy 204 GCCCTTCCCTGGATACCGTCACTTCTTACAAGGTCATTCCCAAAGCAAGTTGCTC-TG 262
Db 121 GCCCTTCCCTGGATACCGTCACTTCTTACAAGGTCATTCCCAAAGCAAGTTGCTC-TG 180
Qy 263 GTGAAGTTGACACCCAGTACCCCTACCGTTCAGAGCAGGATGATTCAGGCTTT-CT 321
Db 181 GTGAAGTTGACACCCAGTACCCCTACCGTTCAGAGCAGGATGATTCAGGCTTTGCT 240
Qy 322 GAAAGCTCGGCTTCAGGAGTATCTCTTGTTGGCAGAGTGGGATCTCAGATTAT-TG 380
Db 241 GAAAGCTCGGCTTCAGGAGTATCTCTTGTTGGCAGAGTGGGATCTCAGATTATGTT 300
Qy 381 GACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGCTACCCA-TC 439
Db 301 GACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGCTACCCAGTC 360
Qy 440 TTCTACCTTTCGGGATGGGACCTTCAGAACCCAGTCCCATACACTGGGGCAGTT-AG 498
Db 361 TTCTACCTTTCGGGATGGGACCTTCAGAACCCAGTCCCATACACTGGGGCAGTTAAG 420
Qy 499 GTTGAGGCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTACCTAGGTATGCTGG-TGC 557
Db 421 GTTGAGGCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTACCTAGGTATGCTGGTTGC 480
Qy 558 CTGCTGTATACGACGCCCTGCGCGGGAGTTTCATCAGGGCTCTGTGTGTGAGG-CCGC 616
Db 481 CTGCTGTATACGACGCCCTGCGCGGGAGTTTCATCAGGGCTCTGTGTGTGAGGCGCCGC 540
Qy 617 CAGGCGCTTTGAAGCAGCGGCAGATTAACCTCTCAAGTGTGAAGGACATCAGAG-AG 675
Db 541 CAGGCGCTTTGAAGCAGCGGCAGATTAACCTCTCAAGTGTGAAGGACATCAGAGGAAG 600
Qy 676 TGGGCGGAGCAATACCTGAAGATCATGGGAGAGATCTTAGACCAAGGGGAGCACTT-CCA 734
Db 601 TGGGCGGAGCAATACCTGAAGATCATGGGAGAGATCTTAGACCAAGGGGAGCACTTCCA 660
Qy 735 GCATCAGAGATGACACCGATCCCAAGGCTGATTGTGAAGAAACAAGATGAGTGAC-GGCAG 793
|||||

Db 661 GCATCAGAGATGACACCGATCCCGAGCTGATGAGAAACAAGATGAGTGACGGGAAG 720
Qy 794 AAGGAGGAGCTCCAGAAGAGCTTAAACATCTGACTGCTTCCAGAAGAGGGGGCCGAG 853
Db 721 AAGGAGGAGCTCCAGAAGAGCTTAAACATCTGACTGCTTCCAGAAGAGGGGGCCGAG 780
Qy 854 AAGGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGTAGGAGGG 913
Db 781 AAGGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGTAGGAGGG 840
Qy 914 GANAGTTAACTGCTGCTGTGANTCCCTGTGGAATATAAGGGGGYMS 962
Db 841 GAGAGTTAACTGCTGCTGTGANTCCCTGTGGAATATAAGGGGGTAG 889
RESULT 7
LOCUS CR541667 786 bp mRNA linear PRI 29-JUN-2004
DEFINITION Homo sapiens full open reading frame cDNA clone RZPD0834E0727D for
gene C12orf8, chromosome 12 open reading frame 8; complete cds,
incl. stopcodon.
ACCESSION CR541667
VERSION CR541667.1 GI:49456294
KEYWORDS Full ORF shuttle clone, Gateway(TM), complete cds.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS Halleck,A., Ebert,L., Mkoundinya,M., Schick,M., Eisenstein,S.,
Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar.W.,
Korn,B., Zuo,D., Hu,Y. and LaBaer,J.
TITLES Cloning of human full open reading frames in Gateway(TM) system
entry vector (pDONR201)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 786)
AUTHORS Halleck,A., Ebert,L., Mkoundinya,M., Schick,M., Eisenstein,S.,
Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar.W.,
Korn,B., Zuo,D., Hu,Y. and LaBaer,J.
TITLES Direct Submision
JOURNAL Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer
Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,
Germany
COMMENT RZPD; RZPD0834E0727D, ORFNO 3284
RZPD; www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=RZPD0834E0727D RZPDLIB;
Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.
834
www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834
www.rzpd.de/products/orfclones/
Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
Contact RZPD (customer.service@rzpd.de) for further information.
This CDS clone is part of a collection of human full ORF clones
jointly established and verified by the Harvard Institute of
Proteomics (HIP) and RZPD.
This CDS has been cloned incl. stopcodon.
The CDS has been inserted into pDONR201 via a BP Clonase(TM)
reaction. Additional sequence has been added in front of the start
codon: att. AAAAAA GCA GGC TCC ACC (ATG).
The stopcodon is followed by the 3' att site: GACCCAGCTTCTT. .att
The clone is validated by full sequence check.
Compared to the reference sequence NM_006817 (GI:13124889) we did
not find any amino acid exchanges.
Clone distribution: http://www.rzpd.de/products/orfclones/.
FEATURES
source Location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="mRNA"

```
/db_xref="taxon:9606"
/clone="RZPD0834E072D"
/clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
/lab_host="DH5Alpha"
/notes="Vector: pDONR201, Site_1: attP1, Site_2: attP2"
1..786
/gene="C12orf8"
CDS
1..786
/gene="C12orf8"
/codon_start=1
/protein_id="CAG46468.1"
/db_xref="GI:49456295"
/translation="MAAAVPRAAFSLPILLGLFLLLSAPHGSGLHTKCALPLDVT
TFYKVIKSKVLVKFDQYPYGKQDFKFLAENSASSDLLVAEIVGISYDGLKLM
ELSEYKLDKESYFVFLFRGDGFENFVYTGAVKVGAIQRWLKGGQVILGPGCLPV
YDALAGEPIRASGVEARQLLKQODNLSSVKETOKKWAQYKIMGKILKQGDFFPA
SEMTRIARLIEKNRMSDGKBELOKSLNLTAFQKGAEKEEL"
ORIGIN
Query Match 52.7%; Score 635.6; DB 9; Length 786;
Best Local Similarity 98.0%; Pred. No. 5.2e-169;
Matches 770; Conservative 0; Mismatches 4; Indels 12; Gaps 12;
Qy 95 ATGGCTGCGCTGTGCCCCGCGCGCATTTCTCTCCCGCTGCTT-CCCTTCTCTCTGGGC 153
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1 ATGGCTGCGCTGTGCCCCGCGCGCATTTCTCTCCCGCTGCTTCTCTCTCTCTCTGGGC 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 154 TTCTGCTCTCTCTCGCTCCGATCGCGCGCGCGCTGCGACCAA-GGCGGCCCTTCCC 212
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 61 TTCTGCTCTCTCTCGCTCCGATCGCGCGCGCGCTGCGACCAAAGGCGCGCTTCCC 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 213 CTGATACGCTCACTTTCTACAGGTCAATTCCTCCAAAGCAAGTTCGTC-TGGTGAAGTTC 271
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 121 CTGATACGCTCACTTTCTACAGGTCAATTCCTCCAAAGCAAGTTCGTC-TGGTGAAGTTC 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 272 GACACCCAGTACCTCTAGGTGAGAGCAGATGATTCAGAGCTTT-CTGAAAACCTCG 330
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 181 GACACCCAGTACCTCTAGGTGAGAGCAGATGATTCAGAGCTTTCTGCTGAAAACCTCG 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 331 GCTTCCAGCATGATCTCTTGGTGCGAGAGTGGGATCTCAGATTAT-GTGCAAGCTG 389
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 241 GCTTCCAGCATGATCTCTTGGTGCGAGAGTGGGATCTCAGATTATGTTGTCAGAGCTG 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 390 AACATGAGCTGAGTGAGAAATACAGCTGACAAAGAGACTACCA-TCTTCTACCTC 448
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 301 AACATGAGCTGAGTGAGAAATACAGCTGACAAAGAGACTACCAAGTCTTCTACCTC 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 449 TTCCGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGTT-AGTTTGGAGCC 507
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 361 TTCCGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGTTAAGTTGGAGCC 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 508 ATCCAGCTGGCTGAAGGGCAAGGGGTCTACCTAGGTATGCTCTGG-TGCTGCTGTA 566
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 421 ATCCAGCTGGTGTGAAGGGCAAGGGGTCTACCTAGGTATGCTCTGGTTGCTGCTGTA 480
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 567 TAGAGCCCTTGGCGGGGAGTTATCAGGCGCTCTGTTGTGAGG-CCGCCAGGCCCTC 625
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 481 TACGATGCTTGGCGGGGAGTTATCAGGCGCTCTGTTGTGAGGCGCCGCGAGCCCTC 540
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 626 TTGAAGCAGGGCGAGATAACTCTCAAGTGTGAGGAGACTCAGAG-AGTGGCCGAG 684
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 541 TTGAAGCAGGGCGAGATAACTCTCAAGTGTGAGGAGACTCAGAGAGAGTGGGGCCGAG 600
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 685 CAATACCTGAAGATCATGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGCATCAGAG 743
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 601 CAATACCTGAAGATCATGGGAAGATCTTAGACCAAGGGGAGCACTTCCAGCATCAGAG 660
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 744 ATGACACGGATGCGCAGCTGATTGAGAAGAACAGATGAGTAC-GGCAGAGGAGAG 802
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 661 ATGACACGGATGCGCAGCTGATTGAGAAGAACAGATGAGTACGCGGAAGAGGAGAG 720
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 803 CTCAGAGAGCTTAACATCTGCTCTTCCAGAGAGGGGGCGAGAAAGAGAG 862
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

FEATURES

```
Db 721 CTCGAGAGAGCTTAACATCCTGACTGCTTCAGAGAAGAGGGCCGCGAGAGAGGAG 780
Qy 863 CTGTAA 868
Db 781 CTGTAA 786

RESULT 8
BC017125 1258 bp mRNA linear ROD 30-JUN-2004
LOCUS Mus musculus endoplasmic reticulum protein 29, mRNA (cdna clone
DEFINITION MGC:27552 IMAGE:4481563), complete cds.
ACCESSION BC017125
VERSION BC017125.1 GI:16877775
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1258)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heish,F.,
Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A.; McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalobon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
GENERATION AND INITIAL ANALYSIS OF MORE THAN 15,000 FULL-LENGTH
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1258)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu,
Parvaneh Saeedi, JR Santos, Angeliq Schnerch, Ursula Skaleka,
Duane Smalls, Jeff Stott, Miranda Tsai, George Yang, Jacque
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 38 Row: a Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19526462.
Location/Qualifiers
```



```
QY 56 CTATCGCTTACCTACCTCCCTCTGAGGAAACCCGGATATGGCTGCGCTGTGCCCCGC 115
Db 5 CTGTAGCTTGCACATCTCCGCTCCACATCCGGCGTGATGGCTGCCGCGCTCTCTGGC 64

QY 116 GCCGATTTCTCTCCCGCTGCT-TCCCTTTCTCTGGGCTTCTGCTCTCTCTCCGCTCCG 174
Db 65 GCGCTTTGCTCTCTCCCACTGCTGTCGTTCTCTCTCTGGCTGCTCTCTCTCTCTCTCT 124

QY 175 CATGGCGGAGCGGCTGACACAA-GGCGCCCTTCCCTGGATACGGTCACTTTCTTAC 233
Db 125 CATGGCGGAGCGGCTGACACAAAGGCGCCCTTCCCTTGGACACAGTCACTTTCTTAC 184

QY 234 AAGGTCATTCCCAAAAGCAAGTTCTGCT-TGGTGAAGTTGACACCAAGTACCCCTACGGT 292
Db 185 AAGGTCATTCCCAAAAGCAAGTTCTGCTTGGTGAAGTTGACACCAAGTACCCCTATGGA 244

QY 293 GAGAACGAGGATGAGTTCAAGCGTCT-TCTGAAAACTCGGCTTCCAGGATGATCTCTTG 351
Db 245 GAGAACGAGGATGAGTTTAAAGCGTCTGGCTGAGAACTCAGCCTCCAGCGATGATCTTTG 304

QY 352 GTGGCAGAGTGGGATCTCAGATTAT-GTGACAAAGCTGAACATGAGAGTGAAGTGAAGA 410
Db 305 GTGGCAGAGTGGGATCTCAGATTATGTCAGAAAGCTGAACATGAGAGTGAAGTGAAGA 364

QY 411 TACAAGCTGGACAAAGAGAGCTACCA-TCTTCTACCTCTTCCGGATGGGACTTTGAG 469
Db 365 TACAAGCTGGACAAAGAGAGCTTACCAGTCTTCTACCTCTTCCGGATGGGACTTTGAG 424

QY 470 AACCCAGTCCCATACACTGGGCGACTT-AGGTTGGAGCCATCCAGCGCTGGCTGAAGGG 528
Db 425 AATCTGTCTCCCATACAGCGGCGACTTAAAGTTGGAGCCATCCAGCGTGGCTCAAGGG 484

QY 529 CAAGGGGTCTACTAGGTATGCTGG-TGCCTGCTGTATAGCAGCGCTTGGCGGGAG 587
Db 485 CAGGGAGTCTATCTGGCATGCTGGATGCTGCTGCTGCTAGTACGATGCTGCGGGCCAG 544

QY 588 TTCATCAGGCTCTGGTGTGAGG-CGCGCAGGCGCTTGAACAGGCGGCAGATAC 646
Db 545 TTCATCAGGCGCTCCAGCAGAGAGCGCCGCGCCATCTGTAACAGCGGCGAGTGGC 604

QY 647 CTCTCAAGTGTGAAGGAGACTCAGAAG-AGTGGGCGGAGCAATACCTGAAGATCATGGG 705
Db 605 CTCTCAGGTGTGAAGGAGACACAGAAGTGGGCGAGTCACTGATCATGATCATGGG 664

QY 706 AAGATCTTAGACCAAGGGAGCACTT-CCAGCATCAGAGATGACACGATCGCCAGGCTG 764
Db 665 AAGATCTTGGACCAAGGTGAAGACTTCCGGGCTCCGAGCTGGCGCGATCAGTAAGCTC 724

QY 765 ATTGAGAAGNACAAGATGATGATGACCGCAGAGGAGGAGCTCCAGAGAGCTTAAACATCC 824
Db 725 ATTGAGAACNAGATGATGAGG--GTAAGAAGGAGAGCTGCAGAGGAGCCTCAACATCC 782

QY 825 TGACTGCTCTCCAGAAAGGGGCGGAGAGAGAGAGCTGTAAAGAGCTGTCTGTGAT 884
Db 783 TCACCGCTTCCGAAAGAGGCGCGAGAGAGAGAGCTGT-----GAGGGACACAGG 836

QY 885 TTTCAGAGGTTTGGTGGGGGTAGGAGGGGAGNAGTTAACTGCTGCTGTGTGANTCCCTTG 944
Db 837 CTGCTAGGTTTGTCTGAGGGGGGAGG--AGAGCGGCTGCTGACTGTGAGGGCCCTGT 894

QY 945 TGAATATTAAGGGGYNKGGGAAAGGAGTAACTAACCCAGATTTCTGAGCGCTGAGTAT 1004
Db 895 GGG--TGGAGGGGGGAGTGGAGCAAGCAGTCTCTGAGGCGAGAGACTGTGCCCCGAG--T 950

QY 1005 GCCTGACATGATGATCAACATGACCATGCTTGGGATGCTCTAG 1049
Db 951 GCCTGACATGATGCTGAGACCATACCTCTGACATCTCTTG 995
```

RESULT 10

RNU36482

LOCUS RNU36482 1139 bp mRNA linear ROD 29-APR-1998

DEFINITION Rattus norvegicus endoplasmic reticulum protein Erp29 precursor,

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 1139)

Mkrtchian,S., Fang,C., Hellman,U. and Ingelman-Sundberg,M.

A stress-inducible rat liver endoplasmic reticulum protein, Erp29

Eur. J. Biochem. 251 (1-2), 304-313 (1998)

98151262

PUBMED 9492298

2 (bases 1 to 1139)

Mkrtchian,S., Fang,C., Hellman,U. and Ingelman-Sundberg,M.

Direct Submission

Submitted (18-SEP-1995) Source: Mktchian, Dept. of Medical

Biochemistry and Biophysics, Karolinska Institute, Doktorsträngen

12B, Stockholm S-171 77, Sweden

On Aug 11, 1997 this sequence version replaced gi:1172121.

COMMENT

FEATURES

source

1. 1139

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/tissue_type="liver endoplasmic reticulum"

44. 826

/note="translation of this region is similar to unknown

protein from 2D-PAGE of liver tissue, Swiss-Prot Accession

Number P30040"

/codon_start=1

/product="endoplasmic reticulum protein Erp29 precursor"

/protein_id="AAC15239.1"

/db_xref="GI:2317800"

/translation="MAAARPGAVSLPILLSVLGLLLSAPHGSLHTKGLPLDVT

TFYKIVPSKFLVLFQDTQPYGKDEFKRLAENSASSDLLLVAEIGISYDGKLNK

ELSKYKLDKESYFVFLFRDGFENPVYSGVAVKGAIRWLQGVYGLQMPGLPA

YDALAGOFIEASSREARQILKQGDGLSGVKETDKWASQYLKIMKILQDGEDFPA

SELARIKLIENKMSKGBELQSLNLTAFRKKGAKEEL"

44. 139

/evidence=experimental

140. 823

/product="endoplasmic reticulum protein Erp29"

812. 823

misc_signal

(note="putative endoplasmic reticulum retrieval signal

(KEEL)"

polya_signal

1096. 1110

ORIGIN

Query Match 41.2%; Score 496.6; DB 10; Length 1139;

Best Local Similarity 78.8%; Pred. No. 1.7e-129;

Matches 792; Conservative 1; Mismatches 187; Indels 25; Gaps 16;

QY 56 CTATCGCTTACCTACCTCCCTCTGAGGAAACCCGGATATGGCTGCGCTGTGCCCCGC 115

Db 5 CTGTAGCTTGCACATCTCCGCTCCACATCCGGCGTGATGGCTGCCGCGTTCCTGGC 64

QY 116 GCGCATTTCTCTCCCGCTGCT-TCCCTTCTCTGGGCTTCTGCTCTCTCCGCTCGG 174

Db 65 GCGCTTTGCTCTCTCCCACTGCTGCTGCTTCTCTGGCCTGCTGCTCTCTGCTCT 124

QY 175 CATGGCGGAGCGGCTGACACAA-GGCGCCCTTCCCTGGATACGGTCACTTTCTTAC 233

Db 125 CATGGCGGAGCGGCTGACACAAAGGCGCCCTTCCCTTGGACACAGTCACTTTCTAC 184

QY 234 AAGGTCATTCCCAAAAGCAAGTTCGTC-TGGTGAAGTTGACACCAAGTACCCCTACGGT 292

Db 185 AAGGTCATTCCCAAAAGCAAGTTCGTC-TGGTGAAGTTGACACCAAGTACCCCTATGGA 244

QY 293 GAGAACGAGGATGATGATCAACGCTCT-TCTGAAAACTGGCTTCCAGGATGATCTCTTG 351


```
Db 245 GAGAAGCAAGATGAGTTTAAAGCGTGTGGCTGAGAACTCAGCCTCCAGCGATGATCTCTTG 304
Qy 352 GTGGCAGAGTGGGATCTCAGATTAT-GTGACAAGCTGAACATGGAGCTGAGTGAGAAA 410
Db 305 GTGGCAGAGTGGGATCTCAGACTATGCTGACAACTGAACATGGAGCTGAGTGAGAG 364
Qy 411 TACAAGCTGGACAAAGAGAGCTACCCA-TCTTCTACCTCTTCCGGGATGGGACTTTTGA 469
Db 365 TACAAGCTGGACAAAGAGAGCTACCACTCTTCTACCTCTTCCGGGATGGGACTTTGAG 424
Qy 470 AACCAGTCCCATACACTGTTGGGCGAGTT-AGTTGGAGCCATCAGCGCTGCTGAAGGG 528
Db 425 AATCCTGTCCCATACAGCGGGGCGAGTTAAAGTTGGAGCCATCCAGCGCTGGCTCAAGGG 484
Qy 529 CAAGGGGTCTACCTAGGTATGCTGG-TGCTGCTGTATACGACGCCCTCGCCGGGGAG 587
Db 485 CAGGAGTCTATCTGGGGCATGCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544
Qy 588 TTCTATCAGGGCTCTGTGTGGAGG-CGCGCAGGCGCTCTTTGAAGCAGGGGCAAGATAAC 646
Db 545 TTCTATCAGGGCTCTCAGCAGAGAGGCGCGCCAGGCCATCTCTGAAACAGGGCGCAGATGGC 604
Qy 647 CTCTCAAGTGTGAAGGAGACTCAGAG-AGTTGGGCGGAGCAATACCTGAAGATCATGGG 705
Db 605 CTCTCAGGTGTGAAGGAGACAGACAAGAAAGTGGGCCAGTCACTACCTGAAGATCATGGG 664
Qy 706 AAGATCTTTAGACCAAGGGGAGCACTT-CCAGCATCAGAGATGACACGGATCGCCAGGCTG 764
Db 665 AAGATCTTTGGACCAAGGTGAAGACTTCCCGGCTCCGAGCTGGCCGGATCAGTAAGCTC 724
Qy 765 ATTGAGAAAGAAAGATGAGTGACGGCAGAGAGGAGCTCCAGAAGAGCTTAAACATCC 824
Db 725 ATTGAGAAAGAAAGATGAGTGAGG--GTAAGAAGGAAGAGCTCAGAGGAGCTCAACATCC 782
Qy 825 TGACTGCTTCCAGAAAGAGGGGCGGAGAAAGAGGAGCTGTAAAGAGCTGTCTGTGAT 884
Db 783 TCACCGCTTCGCAAGAAAGCGCGGAGAGGAGCTGT-----GAGGGCACCCAGG 836
Qy 885 TTTCCAGAGTTTGTGGGGTGAAGGAGGGGAGNAGTTAACTGTGCTGCTGCTGCTGCTGCTG 944
Db 837 CTGCTAGGTTTGTGAGGGCGGGAGG--AGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
Qy 945 TGGAAATATAAGGGGYSKGGGAAAGMGGTACTAAACCCACGATTTCTGAGCCCTGAGTAT 1004
Db 895 GGG--TGGAGGGGCGAGTGGAGCAAGAGCAGTCTTGAGCCAGAGACCTGTGCCCCGAG--T 950
Qy 1005 GCCTGGACATGATGCTAACATGACCATGCTTGGGATGCTCTAG 1049
Db 951 GCCTGGACACTGAGGCTGCTGAGACCATACCTTGGACATCCTTGG 995

RESULT 11
RNERP29PR RNERP29PR 1084 bp mRNA linear ROD 10-MAR-1997
LOCUS R.norvegicus mRNA for Erp29 protein.
DEFINITION Y10264
ACCESSION Y10264
VERSION Y10264.1 GI:1772593
KEYWORDS ER protein; Erp29. (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1
REFERENCE 1
AUTHORS Demmer,J., Zhou,C. and Hubbard,M.J.
TITLE Molecular cloning of Erp29, a novel and widely expressed resident
of the endoplasmic reticulum
JOURNAL FEBS Lett. 402 (2-3), 145-150 (1997)
MEDLINE 97188599
PUBMED 9037184
REFERENCE 2 (bases 1 to 1084)
AUTHORS Hubbard,M.J.
TITLE Direct Submission
```

```
JOURNAL Submitted (24-DEC-1996) M.J. Hubbard, University of Otago,
COMMENT Department of Biochemistry, Box 56 Dunedin, 9001, NEW ZEALAND
FEATURES Related sequences: U36482 and X94910.
source Location/Qualifiers
1..1084
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="dental enamel epithelium (molar)"
/clone_lib="lambda ZAP Express"
/lab_stage="neonate"
/lab_host="E.coli"
15..797
/function="ER protein"
/codon_start=1
/evidence=experimental
/product="Erp29 precursor"
/protein_id="CAA71313.1"
/db_xref="GI:1772594"
/db_xref="GOA:P52555"
/db_xref="UniProt/Swiss-Prot:P52555"
/translation="MAAAVPGAVSLPILLSVLGLLLSAPHGASGLHTKGNLPDITV
TFKVPKSFVLKEDTQPYGKQDEFKRLAENSASSDLLVAEVGISDYGDKLMN
ELSEKYLKDESYFVFLFRDGFENFPYGVKVGAIQRLKQGVYILGPGCLPA
YDALAQFI EASREARQA ILKQDGLSGVKETDKWASOYLKIMKILDOGEDFFA
SELARISKLIENKMSGKBEELQSLNLTAFRKKGAKEEL"
sig_peptide 15..110
/evidence=experimental
mat_peptide 111..794
/product="Erp29"
/evidence=experimental
misc_signal 783..794
/feature="putative ER-retention signal (KEEL)"
polyA_signal 1084..1089
/evidence=not_experimental
ORIGIN
Query Match 39.8%; Score 479.4; DB 10; Length 1084;
Best Local Similarity 78.9%; Pred. No. 1.3e-124;
Matches 771; Conservative 1; Mismatches 179; Indels 26; Gaps 16;
Qy 84 AACCCGCGATATGGCTGCGCGCTGTGCCCCGCGCATTTCTCTCCCGCTGCT-TCCC 142
Db 4 AATCCGCGGTGATGGCTGCGCGCTTCTCTGCGCGCTTCTCTCCCGCATGCTGTCCG 63
Qy 143 TTCTCTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 201
Db 64 TTCTCTGGGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 123
Qy 202 GCGCCCTTCCCTGGATACGGTCACTTTCTACAGGTCACTTCCCAAAAGCAAGTTCCGTC- 260
Db 124 GCGCCCTTCCCTTGGACACAGTCACTTTCTACAGGTCACTTCCCAAAAGCAAGTTCCGTC 183
Qy 261 TGGTGAAGTTGACACACCCAGTACCCCTACCGTGAAGAGCAGGATGAGTTCAAGCGTCT-T 319
Db 184 TGGTGAAGTTGACACACCCAGTACCCCTATGAGAGAGCAAGATGAGTTTAAAGCGTCTGG 243
Qy 320 CTGAAAACCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTAT- 378
Db 244 CTGAGAACTCAGCTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTATG 303
Qy 379 GTGACAACTGAAATGAGCTGAGTGAGAAATACAGCTGACAAAGAGAGAGCTACCCA- 437
Db 304 GTGACAACTGAAATGAGCTGAGTGAGAAATACAGCTGACAAAGAGAGAGCTACCCA 363
Qy 438 TCTTTCTACTTTCGCGGATGGGACTTTGAGAACCCAGTCCCATATACCTGGGCGAGTT- 496
Db 364 TCTTTCTACTTTCGCGGATGGGACTTTGAGAAATCTCTGTCCCATACAGCGGGGAGTTA 423
Qy 497 AGTTGGAGCCATCCAGGCTGCTGAGGGGCAAGGGGTCTACCTAGGTATGCTGTG-T 555
Db 424 AAGTTGGAGCCATCCAGGCTGCTGAGGGGCAAGGGGTCTATCTGTGGGATGCTGTGAT 483
```

```
QY 556 GCCTGCTGTATACGACGCCCTGGCCGGGAGTTTCATCAGGGCCTCTGTGTGGAG-CC 614
Db 484 GTCTGCTCGTACGATGCCCTTGGCGGCCAGTTTCATCGAGGCTCCACGACGAGGCCC 543
QY 615 GCCAGGCCCTCTTGAAGCAGGGGCGAAGATAAACCCTCTCAAGTGTGAAGGAGACTCAGAG- 673
Db 544 GCCAGGCCCTCTGAACACAGGGGCGAGGATGGCCTCTCAGTGTGAAGGAGACAGCAAGA 603
QY 674 AGTGGCCGAGCAGATACCTGAAGATCATGGGGAAGATCTTGAACCAAGGGGAGCATT-C 732
Db 604 AGTGGCCGAGTCACTACCTGAAGATCATGGGGAAGATCTTGAACCAAGGGGAGCATTCC 663
QY 733 CAGCATCAGAGATGACACGATCGCCAGGCTGATTTGAGAGAACAGATGATGACGCA 792
Db 664 CGCCTCCAGCTGGCCCGGATCAGTAAGCTCATTTAGAACAGATGATGAGG--GTAA 721
QY 793 GAAGGAGGAGCTCCAGAGAGCTTAAACATCTCTGACTGCTTCCAGAGAAAGGGGCGCA 852
Db 722 GAAGGAGAGCTGCAGAGAGGCTCAACATCTCACCGCTTCCGCAAGAAAGGGCGCA 781
QY 853 GAAAGAGGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGG 912
Db 782 GAAGGAGGAGCTGTGA-----GGGCACCGGCTGTAGGGTTTGTGAGGGCGGGAGG 834
QY 913 GCANAGTTAACTGCTGGCTGTGANTCCCTCTGGAATATAGGGGGYMSKGGAAAGW 972
Db 835 G--AGAGCGCTGTGACTGTGGGGCCCTGTGGG--TGGAGGGGCGAGTGGAGCAAGC 890
QY 973 GGTACTTAACCCACGATTCCTGAGCCCTGAGTATGCTCTGAGACATTGCTTAACATCACCAT 1032
Db 891 AGTCTTGAGCAGAGACCTGTGCCCGAG--TGCTTGACACTGAGGCTGCTGAGACCAT 948
QY 1033 GCTTGGGATGCTCTAG 1049
Db 949 ACCCTGGACATCCTTGG 965

RESULT 12
Q692426 438 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 37352 from Patent WO02070737.
ACCESSION Q692426
VERSION Q692426.1 GI:42233241
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liew,C.C., Marshall,W.E. and Zhang,H.
TITLES Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 37352 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source
1..438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 33.6%; Score 404.8; DB 6; Length 438;
Best Local Similarity 97.0%; Pred. No. 1.9e-103;
Matches 425; Conservative 5; Mismatches 6; Indels 2; Gaps 2;
QY 691 CTGAGATCATGGGAGAGATCTTAGCAAGGGGAGCACTT-CCAGCATCAGAGATGACA 749
Db 1 CTGAGATCATGGGAGAGATCTTNNCAAGGGGAGGACTTCCAGCATCAGAGATGACA 60
QY 750 CGGATCGCCAGGCTGATTGAGAGAACAGATGAGTGAC-GGCAGAGGAGGAGCTCCAG 808
Db 61 CGGATCGCCAGGCTGATTGAGAGAACAGATGAGTGACGCGGAGAGGAGGAGCTCCAG 120
```

```
QY 809 AAGAGCTTAAACATCCTGACTGCTTCCAGAGAGGGGGCGAGAAAGAGAGAGCTGTAA 868
Db 121 AAGAGCTTAAACATCCTGACTGCTTCCAGAGAGGGGGCGAGAAAGAGAGAGCTGTAA 180
QY 869 AAAGGCTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAGGGGAGTAACTTAACTGCT 928
Db 181 AAAGGCTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAGGGGAGTAACTTAACTGCT 240
QY 929 GGCTGTGANTCCCTTGTGGAATATAGGGGGYMSKGGGAAAGWGTACTTAACCCACGAT 988
Db 241 GGCTGTGAGTCCCTTGTGGAATATAGGGGGGTAGTGGGAAAGTGGTACTTAACCCACGAT 300
QY 989 TCTGAGCCCTCAGATATGCTGACATTTGATCTTAACATGACCATCTTGGGATGTCTTA 1048
Db 301 TCTGAGCCCTCAGATATGCTGACATTTGATCTTAACATGACCATCTTGGGATGTCTTA 360
QY 1049 GCTGTCTGGGGTAGCTGGAGCACTTACTCAGGTGGCTGGTGAATGACACCTCAGAG 1108
Db 361 GCTGTCTGGGGTAGCTGGAGCACTTACTCAGGTGGCTGGTGAATGACACCTCAGAG 420
QY 1109 GAATGAGTGTATAGAGA 1126
Db 421 GAATGAGTGTATAGAGA 438

RESULT 13
G25553 464 bp DNA linear STS 02-JUN-1996
DEFINITION human STS EST47735, sequence tagged site.
ACCESSION G25553
VERSION G25553.1 GI:1347785
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 464)
AUTHORS Hudson,T.
TITLES Whitehead Institute/MIT Center for Genome Research; Physically
JOURNAL Mapped STS
COMMENT Unpublished (1995)
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu
Primer A: GATCGCCAGGCTGATTGAG
Primer B: CACCAAAACCTGGAAAATCA
STS size: 150
PCR Profile:
Preboak:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:
Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: each 4 mM
Tag Polymerase: 0.025 units/ul
Total Vol: 20 ul
Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCL: 10 mM
pH: 9.3
```

Derived from dbEST (genbank accession D25662).

```
FEATURES
  source
    1..464
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /map="536.9 cR from top of Chr12 linkage group"
  STS
    1..150
  primer_bind
    1..19
  primer_bind
    complement(131..150)
  ORIGIN
    Query Match 33.4%; Score 403; DB 11; Length 464;
    Best Local Similarity 96.2%; Pred. No. 6.2e-103;
    Matches 436; Conservative 5; Mismatches 9; Indels 3; Gaps 3;
  Qy 752 GATGCCAGGCTGATTGAGAGACACAGATGAGTACGGC-AGAAGGAGGAGCTCCAGAA 810
  Db 1 GATGCCAGGCTGATTGAGAGACACAGATGAGTACGGC-AGAAGGAGGAGCTCCAGAA 60
  Qy 811 GAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGCCGAGAAAGAGGAGCTGTAAA 870
  Db 61 GAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGCCGAGAAAGAGGAGCTGTAAA 120
  Qy 871 AGGCTGTCTGTGATTTTCCAGGGTTTGTGGGGTGTAGGAGGGGAGATTAACCTGCTGG 930
  Db 121 AGGCTGTCTGTGATTTTCCAGGGTTTGTGGGGTGTAGGAGGGGAGATTAACCTGCTGG 180
  Qy 931 CTGTGANTCCCTTGTGGAATATAGGGGGYMSKGGGAAAGGWTACTTAACCCAGATTC 990
  Db 181 CTGTGAGTCCCTTGTGGAATATAGGGGGTGTAGGAGGGGAGATTAACCCAGATTC 240
  Qy 991 TGAGCCCTGAGTATGCTTGGACATTTGATGCTTAAATGATGCTTGGGATGTCTTAGC 1050
  Db 241 TGAGCCCTGAGTATGCTTGGACATTTGATGCTTAAATGATGCTTGGGATGTCTTAGC 300
  Qy 1051 TGGTCTGGGGATAGCTGGAGCATTACTCAGGTGGCTGGTGAATGACACCTCAGAAGGA 1110
  Db 301 TGGTCTGGGGATAGCTGGAGCATTACTCAGGTGGCTGGTGAATGACACCTCAGAAGGA 360
  Qy 1111 ATGAGTGTCTATAGAGAGA-GAGAGGAGTGTACTGCCCGAGTCTTTGACATGTAATTC 1169
  Db 361 ATNAGTGTCTATAGAGAGA-GAGAGGAGTGTACTGCCCGAGTCTTTGACATGTAATTC 420
  Qy 1170 TCATTCAATAAA-GTTTCAGTGTGTTTGGTTAA 1201
  Db 421 TCATTCAATAAAAGGTTTCAGTGTGTTTGGTTA 453
  RESULT 14
  G27410 464 bp DNA linear STS 28-JUN-1996
  LOCUS human STS SHGC-31784, sequence tagged site.
  DEFINITION G27410
  ACCESSION G27410
  VERSION G27410.1 GI:1396133
  KEYWORDS STS; STS sequence; primer; sequence tagged site.
  SOURCE Homo sapiens (human)
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE 1 (bases 1 to 464)
  AUTHORS Myers R.M.
  JOURNAL Unpublished (1996)
  COMMENT Contact: Richard M. Myers
    Stanford Human Genome Center (SHGC)
    Stanford University School of Medicine
    Department of Genetics, M-344, Stanford, CA 94305, USA
    Tel: 4157259687
    Fax: 4157259689
    Email: myers@shgc.stanford.edu
  Primer A: GATGCCAGGCTGATTGAG
```

Primer B: CACCAAAACCTCGAAAATCA

STS size: 150
PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds

Annealing: 62 degrees C for 23 seconds

Polymerization: 72 degrees C for 30 seconds

PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9600

Protocol:

Template: 25 ng

Primer: each 1 uM

dNTPs: each 200 uM

Taq Polymerase: 0.05 units/ul

Total Vol: 10 ul

Buffer:

MgCl2: 2.5 mM

KCl: 50 mM

Tris-HCl: 20 mM

pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from D25662
-- Washington University/Merck EST sequence.

FEATURES
source

1..464

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/map="19"

1..150

complement(131..150)

ORIGIN

Query Match 33.4%; Score 403; DB 11; Length 464;

Best Local Similarity 96.2%; Pred. No. 6.2e-103;

Matches 436; Conservative 5; Mismatches 9; Indels 3; Gaps 3;

Qy 752 GATGCCAGGCTGATTGAGAGACACAGATGAGTACGGC-AGAAGGAGGAGCTCCAGAA 810

Db 1 GATGCCAGGCTGATTGAGAGACACAGATGAGTACGGC-AGAAGGAGGAGCTCCAGAA 60

Qy 811 GAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGCCGAGAAAGAGGAGCTGTAAA 870

Db 61 GAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGCCGAGAAAGAGGAGCTGTAAA 120

Qy 871 AGGCTGTCTGTGATTTTCCAGGGTTTGTGGGGTGTAGGAGGGGAGATTAACCTGCTGG 930

Db 121 AGGCTGTCTGTGATTTTCCAGGGTTTGTGGGGTGTAGGAGGGGAGATTAACCTGCTGG 180

Qy 931 CTGTGANTCCCTTGTGGAATATAGGGGGYMSKGGGAAAGGWTACTTAACCCAGATTC 990

Db 181 CTGTGAGTCCCTTGTGGAATATAGGGGGTGTAGGAGGGGAGTGTACTTAACCCAGATTC 240

Qy 991 TGAGCCCTGAGTATGCTTGGACATTTGATGCTTAAATGATGCTTGGGATGTCTTAGC 1050

Db 241 TGAGCCCTGAGTATGCTTGGACATTTGATGCTTAAATGATGCTTGGGATGTCTTAGC 300

Qy 1051 TGGTCTGGGGATAGCTGGAGCATTACTCAGGTGGCTGGTGAATGACACCTCAGAAGGA 1110

Db 301 TGGTCTGGGGATAGCTGGAGCATTACTCAGGTGGCTGGTGAATGACACCTCAGAAGGA 360

Qy 1111 ATGAGTGTCTATAGAGAGA-GAGAGGAGTGTACTGCCCGAGTCTTTGACATGTAATTC 1169

Db 361 ATNAGTGTCTATAGAGAGA-GAGAGGAGTGTACTGCCCGAGTCTTTGACATGTAATTC 420

Qy 1170 TCATTCAATAAA-GTTTCAGTGTGTTTGGTTAA 1201

Db 421 TCATTCAATAAAAGGTTTCAGTGTGTTTGGTTA 453

Search completed: July 4, 2005, 01:46:41
Job time : 5116.25 secs

```
RESULT 15
AX887897      AX887897      500 bp      DNA      linear      PAT 18-DEC-2003
LOCUS
DEFINITION   Sequence 3760 from Patent EP1033401.
ACCESSION   AX887897
VERSION     AX887897.1  GI:40046517
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dunas Milne Edwards J.B., Duclert A. and Giordano J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 3760 06-SEP-2000;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1..500
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             70..381
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE93822.1"
                     /db_xref="GI:40046518"
                     /translation="MAAAVPRAAFSLPLLGLLAPHSAPHGSGLHTKALPLDTV
                     TFYKVPKSKFVLVKFDTPYQYGEKQDEFKRLAENSASSDDLLVAEVGISDYGDRTW
                     S"
     sig_peptide     70..129
                     /note="score 6.9 seq AAFLSPLLPLLLG/FL"
ORIGIN
Query Match      33.4%; Score 402.2; DB 6; Length 500;
Best Local Similarity 96.8%; Pred. No. 1.1e-102;
Matches 485; Conservative 0; Mismatches 8; Indels 8; Gaps 7;
QY      26  CACGTGACCGCTGACTCGGGGGTTCTCCACTATCGCTTACTTACCTCCCTCTCGCAGGAA 85
DB      1   CACGTGACCGCTGACTCGGGGGTTCTCCACTATCGCTTACTTACCTCCCTCTCGCAGGAA 60
QY      86  CCCGCGATATGGCTGCGCTGTGCCCGCGCGGATTTCTCCCGCGCTT-CCCTT 144
DB      61  CCCGCGATATGGCTGCGCTGTGCCCGCGCGGATTTCTCTCCCGCTTCCCTT 120
QY      145 CTCTGGGCTTCTGCTCTCTCGCTCCGATCGCGCGCGCTGCACACCAA-GGC 203
DB      121 CTCTGGGCTTCTGCTCTCTCGCTCCGATGCGCGCGCGCTGCACACCAAGGCG 180
QY      204 GCCCTTCCCTGGATACGGTCACTTTCTAAGGTCAATCCCAAGCAAGTTGCTC-TG 262
DB      181 GCCCTTCCCTGGATACGGTCACTTTCTAAGGTCAATCCCAAGCAAGTTGCTCTTG 240
QY      263 GTGAAGTTTCGACACCCAGTACCCTACGCTGAGAGCAGGATGATTCAGCGTCTT-CT 321
DB      241 GTGAAGTTTCGACACCCAGTACCCTACGCTGAGAGCAGGATGATTCAGCGTCTTGTCT 300
QY      322 GAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTATGTG 381
DB      301 GAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTATGTG 360
QY      382 ACAAGCTGAAATGAGCTGAGTGAGAAATACAGCTGGAACAAGAGAGCTACCCA-TCCT 440
DB      361 GACAAGCGAATACGAGCTGAGTGAGAAATACAGCTGGAACAAGAGAGCTACCCAGTCT 420
QY      441 TCTACTCTTCGGGATGGGACTTTGAGACCCAGTCCCATACACTGGGGCAGTT--AG 498
DB      421 TCTACTCTTCGGGATGGGACTTTGAGAA-CCAGTCCCATACACTGGGGCAGTTAAAG 479
QY      499 GTTGAGCCATCCAGCGCTGG 519
DB      480 GTTGAGCCATCCAGCGCTGG 500
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 05:45:02 ; Search time 615.67 Seconds

(without alignments)
11586.223 Million cell updates/sec

Title: US-08-731-499-7

Perfect score: 1205

Sequence: 1 GCGCGCGAGTCCGCCCCC.....TCAGTGTTCGTTAAGTGG 1205

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-----------|-------------|
| 1 | 1199.4 | 99.5 | 1205 | AAV09021 | Homo sapi |
| 2 | 1062.8 | 88.2 | 1334 | AACT78172 | Human can |
| 3 | 980 | 81.3 | 1600 | AAI59481 | Human pol |
| 4 | 857.4 | 71.2 | 1242 | AA578847 | DNA encod |
| 5 | 803.8 | 66.7 | 963 | ADB75219 | Prostate |
| 6 | 803.8 | 66.7 | 963 | ACN39561 | Tumour-as |
| 7 | 638.4 | 53.0 | 1018 | AA583795 | DNA encod |
| 8 | 517.2 | 42.9 | 4529 | ADBS5829 | Toxicity- |
| 9 | 517.2 | 42.9 | 4529 | ADBS5348 | Primary r |
| 10 | 517.2 | 42.9 | 4529 | ABT42350 | Toxicity |
| 11 | 496.6 | 41.2 | 1139 | ADP72860 | Renal tox |
| 12 | 403 | 33.4 | 464 | AA22418 | Human gen |
| 13 | 402.2 | 33.4 | 500 | AA03762 | Human sec |
| 14 | 369 | 30.6 | 2087 | AA583797 | DNA encod |
| 15 | 328.8 | 27.3 | 395 | ADT95875 | Colon can |
| 16 | 321 | 26.6 | 405 | ACH18638 | Human adu |
| 17 | 312.8 | 26.0 | 629 | ADQ051247 | Novel can |
| 18 | 279.2 | 23.2 | 450 | AAI12456 | Probe #23 |
| 19 | 279.2 | 23.2 | 450 | ABA54164 | Human foe |
| 20 | 279.2 | 23.2 | 450 | AAI33810 | Probe #24 |

| | | | | | | |
|----|-------|------|-----|---|----------|--------------------|
| 21 | 279.2 | 23.2 | 450 | 4 | ABA43707 | Aba43707 Human bre |
| 22 | 279.2 | 23.2 | 450 | 4 | ABA23909 | Aba23909 Probe #23 |
| 23 | 279.2 | 23.2 | 450 | 4 | AAK27875 | Aak27875 Human bon |
| 24 | 279.2 | 23.2 | 450 | 4 | AAK02434 | Aak02434 Human bra |
| 25 | 279.2 | 23.2 | 450 | 4 | ABS27459 | Abs27459 Human liv |
| 26 | 279.2 | 23.2 | 450 | 5 | AAI02367 | AAI02367 Probe #23 |
| 27 | 279.2 | 23.2 | 450 | 6 | ABS02335 | Abs02335 Human gen |
| 28 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 29 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 30 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 31 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 32 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 33 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 34 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 35 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 36 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 37 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 38 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 39 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 40 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 41 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 42 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 43 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 44 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |
| 45 | 279.2 | 23.2 | 450 | 9 | ACH31743 | Ach31743 Human bon |

ALIGNMENTS

RESULT 1
AAV09021
ID AAV09021 standard; cDNA; 1205 BP.
XX AC AAV09021;
XX DT 21-JUL-1998 (first entry)
XX DE Homo sapiens 20ql3 amplicon lb4 transcript.
XX KW 20ql3 amplicon; chromosome 20; tumour; detection;
XX KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
XX KW treatment; age-related macular degeneration; retinitis pigmentata;
XX KW Leber's congenital amaurosis; serine threonine kinase; ds.
XX OS Homo sapiens.
XX PN WO9802539-A1.
XX PD 22-JAN-1998.
XX PF 15-JUL-1997; 97WO-US012343.
XX PR 15-JUL-1996; 96US-00680395.
XX PR 16-OCT-1996; 96US-00731499.
XX PR 17-JAN-1997; 97US-00785532.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX WPI; 1998-110587/10.
XX DR New sequences from the 20ql3 amplicon - used for detecting chromosomal
XX PT abnormalities, particularly tumours, and for developing products for
XX PT treating diseases.
XX PS Claim 1; Page 63; 91pp; English.
XX CC The sequence is that of a cDNA sequence lb4 which was isolated from the
XX CC 20ql3 amplicon. It encodes a serine threonine kinase and can be used as a
XX CC probe for the detection of chromosomal abnormalities at 20ql3. It and
XX CC other sequences isolated from the 20ql3 amplicon are consistently

CC amplified in primary tumours. These sequences are useful as probes or as
CC probe targets for monitoring the relative copy number of corresponding
CC sequences from a biological sample such as tumour cells. The sequences
CC can also be used in therapeutic applications for modulating the
CC expression of the endogenous gene or the activity of the gene product.
CC Examples of therapeutic approaches include antisense inhibition of gene
CC expression, gene therapy, and monoclonal antibodies that specifically
CC bind the gene products. The products can also be used in the treatment of
CC other diseases, e.g. age-related macular degeneration, Leber's congenital
CC amaurosis and retinitis pigmentata

XX Sequence 1205 BP; 277 A; 298 C; 362 G; 261 T; 0 U; 7 Other;

| | | | | | |
|-----------------------|-----|-----------------|---|---------------|--------------|
| Query Match | | 99.5%; | Score 1199.4; | DB 2; | Length 1205; |
| Best Local Similarity | | 99.9%; | Pred. No. 0; | | |
| Matches 1204; | | Conservative 0; | Mismatches 1; | Indels 0; | Gaps 0; |
| Qy | 1 | CGCGCGTGA | TCGCGCCCCCGACGTCACCGCTGACCTCGGGGGTTC | CCACTATC | 60 |
| Db | 1 | CGCGCGTGA | TCGCGCCCCCGACGTCACCGCTGACCTCGGGGGTTC | CCACTATC | 60 |
| Qy | 61 | GCTTACCTAC | CTCTCGAGGAACCGCGATATGCTCGCGCTGTCGCCCGCGC | 120 | |
| Db | 61 | GCTTACCTAC | CTCTCGAGGAACCGCGATATGCTCGCGCTGTCGCCCGCGC | 120 | |
| Qy | 121 | ATTTCTCTCC | CGCTCTTCTCTCGGGCTTCCTGCTCCTCTCGCTCCGCTCCG | CATGGC | 180 |
| Db | 121 | ATTTCTCTCC | CGCTCTTCTCTCGGGCTTCCTGCTCCTCTCGCTCCGCTCCG | CATGGC | 180 |
| Qy | 181 | GGCAGCGCCT | GCACACAAAGCGCCCTTCCTGGATACCGGTCACTTCTACA | AGGTCA | 240 |
| Db | 181 | GGCAGCGCCT | GCACACAAAGCGCCCTTCCTGGATACCGGTCACTTCTACA | AGGTCA | 240 |
| Qy | 241 | TTCCCAAAG | CAAGTTCGTCGTAAGTTCGACACCCAGTACCCCTACG | GTGAAGCA | 300 |
| Db | 241 | TTCCCAAAG | CAAGTTCGTCGTAAGTTCGACACCCAGTACCCCTACG | GTGAAGCA | 300 |
| Qy | 301 | GGATGAGTT | CAAGCGTCTTCTGMAAATCGGCTCCAGCGATGATCTCT | TGGTGGCAG | 360 |
| Db | 301 | GGATGAGTT | CAAGCGTCTTCTGMAAATCGGCTCCAGCGATGATCTCT | TGGTGGCAG | 360 |
| Qy | 361 | GTGGGATCT | CAGATTATGTGACAACTGAACTGAGCTGAGTGAAGAA | TACAAGCTGG | 420 |
| Db | 361 | GTGGGATCT | CAGATTATGTGACAACTGAACTGAGCTGAGTGAAGAA | TACAAGCTGG | 420 |
| Qy | 421 | ACAAAGAG | CTACCACTTCTACCTCTTCGGGATGGGACTTTGAGAC | CCAGTCCC | 480 |
| Db | 421 | ACAAAGAG | CTACCACTTCTACCTCTTCGGGATGGGACTTTGAGAC | CCAGTCCC | 480 |
| Qy | 481 | ATACACTGG | GGCAGTTAGTGGAGCCATCCAGCGCTGGCTGAAGGG | CAAGGGTCTAC | 540 |
| Db | 481 | ATACACTGG | GGCAGTTAGTGGAGCCATCCAGCGCTGGCTGAAGGG | CAAGGGTCTAC | 540 |
| Qy | 541 | CTAGGTATG | CTGTGCTGCTGATACGACCGCTCGCGGGGAGTTCA | TCAAGGGCCT | 600 |
| Db | 541 | CTAGGTATG | CTGTGCTGCTGATACGACCGCTCGCGGGGAGTTCA | TCAAGGGCCT | 600 |
| Qy | 601 | CTGGTGTGG | GGCCCGCCGCTTCTGAACGAGGGCAAGTACCTCTCA | AGTGTGAA | 660 |
| Db | 601 | CTGGTGTGG | GGCCCGCCGCTTCTGAACGAGGGCAAGTACCTCTCA | AGTGTGAA | 660 |
| Qy | 661 | GGAGACTCA | GAGTGGCGCGAGCAATACCTGAAGATCATGGGAAG | ATCTTAGACCA | 720 |
| Db | 661 | GGAGACTCA | GAGTGGCGCGAGCAATACCTGAAGATCATGGGAAG | ATCTTAGACCA | 720 |
| Qy | 721 | GGGGAGCA | CTTCCAGCATCAGAGATGACACCGATCCCGAGCTG | ATTGAGAAGACA | 780 |
| Db | 721 | GGGGAGCA | CTTCCAGCATCAGAGATGACACCGATCCCGAGCTG | ATTGAGAAGACA | 780 |
| Qy | 781 | TCAGTGA | CGGCAAGAGGAGCTCCAGAGAGCTTAAACATCTC | TGACTGCTTCCAG | 840 |
| Db | 781 | TCAGTGA | CGGCAAGAGGAGCTCCAGAGAGCTTAAACATCTC | TGACTGCTTCCAG | 840 |

| | | | | |
|----------|---|-------------|---|------|
| Qy | 841 | GAAGGGGGCCG | AGAAAGAGGAGCTGTAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTG | 900 |
| Db | 841 | GAAGGGGGCCG | AGAAAGAGGAGCTGTAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTG | 900 |
| Qy | 901 | GGGGTAGGAGG | AGANAGTTTAACTCTGCTGTGANTCCCTTGTGGAATATAAGGGGGY | 960 |
| Db | 901 | GGGGTAGGAGG | AGANAGTTTAACTCTGCTGTGANTCCCTTGTGGAATATAAGGGGGY | 960 |
| Qy | 961 | MSKGGGAAA | AGWGCTACTAAACCCACGATTCTGAGCCCTGAGTATGCCCTGGACATTGATGC | 1020 |
| Db | 961 | MSKGGGAAA | AGWGCTACTAAACCCACGATTCTGAGCCCTGAGTATGCCCTGGACATTGATGC | 1020 |
| Qy | 1021 | TAACATGACCA | TGCTTGGGATGTCTCTAGCTGTCTGGGATAGCTGGAGCATTACTCA | 1080 |
| Db | 1021 | TAACATGACCA | TGCTTGGGATGTCTCTAGCTGTCTGGGATAGCTGGAGCATTACTCA | 1080 |
| Qy | 1081 | GGTGGCTGTG | TAAATGACACCTCAGAAGGAATCAGTCTATAGAGAGGAGGAGTGT | 1140 |
| Db | 1081 | GGTGGCTGTG | TAAATGACACCTCAGAAGGAATCAGTCTATAGAGAGGAGGAGTGT | 1140 |
| Qy | 1141 | ACTGCCCAGGT | CTTTGACAGATGTAATTTCTCAATTAAAGTTTCAGTGTTTTGGTTA | 1200 |
| Db | 1141 | ACTGCCCAGGT | CTTTGACAGATGTAATTTCTCAATTAAAGTTTCAGTGTTTTGGTTA | 1200 |
| Qy | 1201 | AGTGG | 1205 | |
| Db | 1201 | AGTGG | 1205 | |
| RESULT 2 | | | | |
| AAC78172 | | | | |
| ID | AAC78172 standard; cDNA; 1334 BP. | | | |
| XX | AAC78172; | | | |
| AC | AAC78172; | | | |
| XX | 08-FEB-2001 (first entry) | | | |
| DT | Human cancer associated gene sequence SEQ ID NO:566. | | | |
| XX | Human; cancer associated gene; cancer antigen; detection; cancer; | | | |
| KW | diagnosis; cytostatic; proliferative; vulnerable; immunomodulator; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| KW | antidiabetic; antistatic; antirheumatic; antidiabetic; antitumor; | | | |
| OS | Homo sapiens. | | | |
| XX | WO200055350-A1. | | | |
| PN | 21-SEP-2000. | | | |
| XX | 08-MAR-2000; 2000WO-US005882. | | | |
| PF | 12-MAR-1999; 99US-0124270P. | | | |
| XX | (HUMA-) HUMAN GENOME SCI INC. | | | |
| XX | Rosen CA, Ruben SM; | | | |
| PI | WPI; 2000-587533/55. | | | |
| XX | P-PSDB; AAB43963. | | | |
| DR | Novel isolated nucleic acids comprising sequences encoding peptides | | | |
| XX | useful for treating or diagnosing e.g. cancer. | | | |
| PT | Claim 1; Page 1091; 2352pp; English. | | | |
| XX | AAC77607 to AAC78448 encode the human cancer associated proteins given in | | | |
| PS | | | | |
| XX | | | | |
| CC | | | | |

CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antiadhesive; antiaethmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatologic; neuroprotective; cardiatic; thrombolytic; coagulant;
 CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention

XX
 SQ Sequence 1334 BP; 337 A; 319 C; 396 G; 278 T; 0 U; 4 Other;

Query Match 88.2%; Score 1062.8; DB 3; Length 1334;

Best Local Similarity 98.2%; Pred. No. 6.5e-307;

Matches 1193; Conservative 5; Mismatches 5; Indels 12; Gaps 12;

QY 1 GCGCGGTGAGTCCGCCCCCAGTACGTCACCGCTGACTCGGGCGTCTCCCATATC 60
 DB 40 GCGCGGTGAGTCCGCCCCCAGTACGTCACCGCTGACTCGGGCGTCTCCCATATC 99
 QY 61 GCTTACCTACCTCCCTCTGCGAGAACCGCGGATATGCTGCGCTGCGCCCGCGCGC 120
 DB 100 GCTTACCTACCTCCCTCTGCGAGAACCGCGGATATGCTGCGCTGCGCCCGCGCGC 159
 QY 121 ATTCTCTCCCGCTGCTT-CCCTTCTCTCGGCTTCTGCTCTCTCGCTCGCATGG 179
 DB 160 ATTCTCTCCCGCTGCTTCCCTTCTCTCGGCTTCTGCTCTCTCGCTCGCATGG 219
 QY 180 CGCAGCGCGCTGCACACAA-GGCGCCCTTCCCTGGATACGGTCACTTCTACAAGGT 238
 DB 220 CGCAGCGCGCTGCACACAAAGGCGCCCTTCCCTGGATACGGTCACTTCTACAAGGT 279
 QY 239 CATTCCTCAAGAGATTCGTC-TGGTGAAGTTGACACCCAGTACCCCTACGGTGAGAA 297
 DB 280 CATTCCTCAAGAGATTCGTC-TGGTGAAGTTGACACCCAGTACCCCTACGGTGAGAA 339
 QY 298 CGAGATGAGTTCAAGCGTCTT-CTGAAACTCGGCTTCCAGCGATGATCTCTTGGTGGC 356
 DB 340 CGAGATGAGTTCAAGCGTCTTCTGAAACTCGGCTTCCAGCGATGATCTCTTGGTGGC 399
 QY 357 AGAGTGGGGATCTCAGATTAT-GTGACAAGCTGAACATGGAGCTGAGTGAGAAATACAA 415
 DB 400 AGAGTGGGGATCTCAGATTATGTTGACAGCTGAACATGGAGCTGAGTGAGAAATACAA 459
 QY 416 GCTGACAAAGAGAGCTACCA-TCTTCTACCTCTTCCGGATGGGACTTTGAGAACCC 474
 DB 460 GCTGACAAAGAGAGCTACCACTCTTACCTCTTCCGGATGGGACTTTGAGAACCC 519
 QY 475 AGTCCATACACTGGGCGAGTT-AGGTTGGAGCCATCCAGCGCTGGCTGAGAGGGCGAAGG 533
 DB 520 AGTCCATACACTGGGCGAGTTAAGTTGGAGCCATCCAGCGCTGGCTGAGAGGGCGAAGG 579
 QY 534 GGTCTACCTAGTATGCTGG-TGCTGCTGTATACGACGCCCTTGGCGGGAGTTCTAT 592
 DB 580 GGTCTACCTAGTATGCTGGTGTGCTGCTGTATACGACGCCCTTGGCGGGAGTTCTAT 639
 QY 593 CAGGCGCTCTGTTGGAGG-CCGCCAGCCCTCTTGAAGCAGGGCGGCAAGATAACCTCTC 651
 DB 640 CAGGCGCTCTGTTGGAGGCGCGCCCTCTTGAAGCAGGGCGGCAAGATAACCTCTC 699
 QY 652 AAGTGTGAAGGAGCTCAGAAG-AGTGGCGCGAGCAATACCTGAAGATCATGGGGAAGAT 710

DB 700 AAGTGTGAAGGAGACTCAGAAAGTGGCCGCGCAANTACCTGAAGATCATGGGGAAGAT 759
 QY 711 CTTAGACAAGGGGAGGACATT-CCAGATCAGAGATGACACGGATCGCCAGGCTGATTGA 769
 DB 760 CTTAGACAAGGGGAGGACTTTCCAGCATCAGAGATGACACGGATCGCCAGGCTGATTGA 819
 QY 770 GAACAACAAGATGATGAC-GGCAGAAGGAGGAGCTCCAGAAGAGCTTAAACATCCTGAC 828
 DB 820 GAACAACAAGATGATGACGCGGAAGAAGGAGGAGCTCCAGAAGAGCTTAAACATCCTGAC 879
 QY 829 TGCCTTCCAGAAGAGGGGCGCGAGAAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTC 888
 DB 880 TGCCTTCCAGAAGAGGGGCGCGAGAAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTC 939
 QY 889 CAGGTTTGGTGGGGTAGGGAGGGAGGAGTAACTTCTGCTGGCTGTGANTCCCTTGTGGA 948
 DB 940 CAGGTTTGGTGGGGTAGGGAGGGAGGAGTAACTTCTGCTGGCTGTGANTCCCTTGTGGA 999
 QY 949 ATATAAGGGGYSKSGGAAAGGAGTAACTTCTGAGGCTTCTGAGGCTGAGTATGCT 1008
 DB 1000 ATATAAGGGGAGTGGGAAAGTGGTAACTTCTGAGGCTTCTGAGGCTGAGTATGCT 1059
 QY 1009 GGACATTGATCTAACATGACCATGCTTGGGATGTCTTAGCTGTCTGGGGATAGCTGG 1068
 DB 1060 GGACATTGATCTAACATGACCATGCTTGGGATGTCTTAGCTGTCTGGGGATAGCTGG 1119
 QY 1069 AGCACTTACTCAGTGGCTGGTGAATGACACCTCAGAAGAAATGAGTGTCTATAGAGG 1128
 DB 1120 AGCACTTACTCAGTGGCTGGTGAATGACACCTCAGAAGAAATGAGTGTCTATAGAGG 1179
 QY 1129 AGAGAGGAGTACTGCCAGGCTTTTGACAGATGTAATCTTCAATCAATTAAGTTTCA 1188
 DB 1180 AGAGAGGAGTACTGCCAGGCTTTTGACAGATGTAATCTTCAATCAATTAAGTTTCA 1239
 QY 1189 GTGTTTGGTTAACT 1203
 DB 1240 GTGTTTGGTTAANT 1254

RESULT 3
 AA159481
 ID AA159481 standard; cDNA; 1600 BP.
 XX AA159481;
 AC AA159481;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 1684.
 XX
 KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-0052317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.

| | | |
|----|--|--|
| XX | (HYSE-) HYSEQ INC. | |
| XX | Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; | |
| XX | PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA; | |
| XX | PI Zhou P, Goodrich R, Drmanac RT; | |
| XX | WPI; 2001-442253/47. | |
| XX | DR P-PSDB; AAM40325. | |
| XX | Novel nucleic acids and polypeptides, useful for treating disorders such | |
| XX | PT as central nervous system injuries. | |
| XX | Claim 1; SEQ ID NO 1684; 10078pp; English. | |
| XX | The invention relates to human nucleic acids (AAI57798-AAI61369) and the | |
| XX | CC encoded polypeptides (AAM38642-AAM42213) with nootropic, | |
| XX | CC immunosuppressant and cytoskeleton activity. The polynucleotides are useful | |
| XX | CC in gene therapy. A composition containing a polypeptide or polynucleotide | |
| XX | CC of the invention may be used to treat diseases of the peripheral nervous | |
| XX | CC system, such as peripheral nervous injuries, peripheral neuropathy and | |
| XX | CC localised neuropathies and central nervous system diseases, such as | |
| XX | CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic | |
| XX | CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the | |
| XX | CC utilisation of the activities such as: Immune system suppression, | |
| XX | CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic | |
| XX | CC and thrombolytic activity, cancer diagnosis and therapy, drug screening, | |
| XX | CC assays for receptor activity, arthritis and inflammation, leukaemias and | |
| XX | CC C.N.S disorders. Note: The sequence data for this patent did not form | |
| XX | CC part of the printed specification | |
| XX | SEQ Sequence 1600 BP; 402 A; 353 C; 431 G; 414 T; 0 U; 0 Other; | |
| XX | Query Match 81.3%; Score 980; DB 4; Length 1600; | |
| XX | Best Local Similarity 97.9%; Pred. No. 4.3e-282; | |
| XX | Matches 1112; Conservative 5; Mismatches 7; Indels 12; Gaps 12; | |
| Qy | 82 GGAACCCCGCGATATGGTCGGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTT-C 140 | |
| Db | 80 GAATTCGCGCATATGGTCGGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTTC 139 | |
| Qy | 141 CTTCTCTCGGGCTTCTGTCTCTCTCGCTCCGATGGCGGCGCGCTGCGACACCAA 200 | |
| Db | 140 CTTCTCTCGGGCTTCTGTCTCTCTCGCTCCGATGGCGGCGCGCTGCGACACCAA 199 | |
| Qy | 201 -GGCGCCCTCCCGTGGATAGCTGCTTCTGAGAGTCAATCCCAAGGTCATTC 259 | |
| Db | 200 GGGCGCCCTCCCGTGGATAGCTGCTTCTGAGAGTCAATCCCAAGGTCATTC 259 | |
| Qy | 260 C-TGGTGAAGTTCCGACACCCAGTACCCCTACGGTGAGAGGAGGATGAGTTCAAGCGTCT 318 | |
| Db | 260 CTGGTGAAGTTCCGACACCCAGTACCCCTACGGTGAGAGGAGGATGAGTTCAAGCGTCT 319 | |
| Qy | 319 T-CTGAAACCTCGGCTTCACGGATGATCTCTTGGTGGCAGAGTGGGGATCTCAGATTA 377 | |
| Db | 320 TGCTGAAACCTCGGCTTCACGGATGATCTCTTGGTGGCAGAGTGGGGATCTCAGATTA 379 | |
| Qy | 378 T-GTGACAGCTGACATGGAGCTGAGTGAGAAATACAGCTGGACAGAGAGCTACCC 436 | |
| Db | 380 TGGTGACAGCTGACATGGAGCTGAGTGAGAAATACAGCTGGACAGAGAGCTACCC 439 | |
| Qy | 437 A-TCTTCTACCTTTCGGGATGGGACTTTTGAGAACCCAGTCCCATACACTGGGGCAGT 495 | |
| Db | 440 AGTCTTCTACCTTTCGGGATGGGACTTTTGAGAACCCAGTCCCATACACTGGGGCAGT 499 | |
| Qy | 496 T-AGGTTGGAGCATTCACGCGCTGGCTGAAAGGGGCAAGGGTCTACCTAGTATGCTCTGG 554 | |
| Db | 500 TAAGGTTGGAGCATTCACGCGCTGGCTGAAAGGGGCAAGGGTCTACCTAGTATGCTCTGG 559 | |
| Qy | 555 -TGCTGCTGTATACGACGCCCTGGCGCGGGAGTTTCATCAGGGCTCTGTGTGGAGG- 612 | |
| Db | 560 TTGCTGCTGTATACGACGCCCTGGCGCGGGAGTTTCATCAGGGCTCTGTGTGGAGG- 619 | |
| Qy | 613 CCGCCAGGCCCTCTTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAA 672 | |
| Db | 620 CCGCCAGGCCCTCTTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAA 679 | |
| Qy | 673 G-AGTGGGCGGACCAATACCTGAGATCATGGGGAGATCTTAGACCAAGGGGAGCACTT 731 | |
| Db | 680 GAAAGTGGGCGGACCAATACCTGAGATCATGGGGAGATCTTAGACCAAGGGGAGCACTT 739 | |
| Qy | 732 -CCAGCATCAGAGATCAGCGGATCGCCAGGCTGATTGAGAAGAAACAAGATGAGTGAC-G 789 | |
| Db | 740 CCAGCATCAGAGATCAGCGGATCGCCAGGCTGATTGAGAAGAAACAAGATGAGTGACGG 799 | |
| Qy | 790 GCAGAAGGAGAGCTCCAGAAGAGCTTAAACATCCTGACTGCTCCAGAAAGAGGGGGC 849 | |
| Db | 800 GAAGAAGGAGAGCTCCAGAAGAGCTTAAACATCCTGACTGCTCCAGAAGAGGGGGC 859 | |
| Qy | 850 CGAGAAAGAGAGCTGTAAAGAGGCTGTCTGTGATTTTCCAGGGTTTGTGGGGTAGGG 909 | |
| Db | 860 CGAGAAAGAGAGCTGTAAAGAGGCTGTCTGTGATTTTCCAGGGTTTGTGGGGTAGGG 919 | |
| Qy | 910 AGGGGAAAGTTAACTGCTGCTGTGANTCCCTTGTGGAATATAAGGGGGMKGGGAAA 969 | |
| Db | 920 AGGGGAGAGTTAACTGCTGCTGTGAGTCCCTTGTGGAATATAAGGGGAGTAGGGAAA 979 | |
| Qy | 970 AGMGGTACTAACCCACGATTTCTGAGCCCTGAGTATGCTTGACATGATGATTAACATGAC 1029 | |
| Db | 980 AGTGGTACTAACCCACGATTTCTGAGCCCTGAGTATGCTTGACATGATGATTAACATGAC 1039 | |
| Qy | 1030 CATGCTTTGGGATGCTCTAGCTGCTGGGGATAGCTGGAGCACTTACTCAGGTGGCTGG 1089 | |
| Db | 1040 CATGCTTTGGGATGCTCTAGCTGCTGGGGATAGCTGGAGCACTTACTCAGGTGGCTGG 1099 | |
| Qy | 1090 TGAATGACACCTCAGAAAGGAATGAGTGTCTATAGAGAGGAGAGAGTGTACTGCCCCAG 1149 | |
| Db | 1100 TGAATGACACCTCAGAAAGGAATGAGTGTCTATAGAGAGGAGAGAGTGTACTGCCCCAG 1159 | |
| Qy | 1150 GTCTTTGACAGATGAATTTCTCATTTCAATTAAGTTTCAGTGTGTTTGGTTAAGTGG 1205 | |
| Db | 1160 GTCTTTGACAGATGAATTTCTCATTTCAATTAAGTTTCAGTGTGTTTGGTTAAGTGG 1215 | |
| XX | RESULT 4 | |
| XX | AAS78847 | |
| ID | AAS78847 standard; cDNA; 1242 BP. | |
| XX | AC AAS78847; | |
| XX | XX | |
| DT | 13-FEB-2002 (first entry) | |
| XX | DNA encoding novel human diagnostic protein #14651. | |
| DE | Human; chromosome mapping; gene mapping; gene therapy; forensic; | |
| KW | food supplement; medical imaging; diagnostic; genetic disorder; ss. | |
| XX | Homo sapiens. | |
| OS | WO200175067-A2. | |
| XX | XX | |
| PN | 11-OCT-2001. | |
| XX | XX | |
| PD | 30-MAR-2001; 2001WO-US008631. | |
| PF | 31-MAR-2000; 2000US-00540217. | |
| XX | XX | |
| PR | 23-AUG-2000; 2000US-00649167. | |
| XX | XX | |
| XX | (HYSE-) HYSEQ INC. | |
| PA | Drmanac RT, Liu C, Tang YT; | |
| PI | WPI; 2001-639362/73. | |
| XX | DR P-PSDB; ABG14660. | |
| XX | New isolated polynucleotide and encoded polypeptides, useful in | |

XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
PI Hoersh S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;
XX WPI; 2003-248033/24.
XX New nucleic acid molecule, useful for diagnosing or treating prostate
PT cancer.
XX Disclosure; SEQ ID NO 43; 99pp; English.
PS
CC The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.
CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 963 BP; 225 A; 240 C; 297 G; 201 T; 0 U; 0 Other;

Query Match 66.7%; Score 803.8; DB 10; Length 963;
Best Local Similarity 97.8%; Pred. No. 1.8e-229;
Matches 934; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

Qy 87 CCGCGGATATGGCTGCGCGCTGTCGCCGCGCGCATTTCTTCCCGCTGTT-CCCTTC 145
Db 4 CCGCGGATATGGCTGCGCGCTGTCGCCGCGCGCATTTCTTCCCGCTGTTCCCTTC 63
Qy 146 TCCTTGGGCTTCTGCTCTCTCCGCTCGCATGGCGGCGGCTGCACACCAA-GGCG 204
Db 64 TCCTTGGGCTTCTGCTCTCTCCGCTCGCATGGCGGCGGCTGCACACCAAGGGCG 123
Qy 205 CCCTTCCCTCGATACCGGTCACTTTCTACAAAGGTCAITPCCAAAGCAAGTTCGTC-TGG 263
Db 124 CCCTTCCCTCGATACCGGTCACTTTCTACAAAGGTCAITPCCAAAGCAAGTTCGTC-TGG 183
Qy 264 TGAAGTTGACACCCAGTACCCCTACGCTGAGAGCAGGATGAGTTCAGCGCTT-CTG 322
Db 184 TGAAGTTGACACCCAGTACCCCTACGCTGAGAGCAGGATGAGTTCAGCGCTTCTGCTG 243
Qy 323 AAAAATCTCGGCTTCCAGCGATGATCTTGGTGGCAGAGGTGGGATCTCAGATTAT-GTG 381
Db 244 AAAATCTCGGCTTCCAGCGATGATCTTGGTGGCAGAGGTGGGATCTCAGATTATGCTG 303
Qy 382 ACAAGCTGAAATGAGGTGAGTGAAGAATAACAAGCTGGACAAAGAGAGCTACCGAGTCT 440
Db 304 ACAAGCTGAAATGAGGTGAGTGAAGAATAACAAGCTGGACAAAGAGAGCTACCGAGTCT 363
Qy 441 TCTACCTTTCCGGATGGGACTTTGAGACCCAGTCCCATACACTGGGCGAGTT-AGG 499
Db 364 TCTACCTTTCCGGATGGGACTTTGAGACCCAGTCCCATACACTGGGCGAGTTAAGG 423
Qy 500 TTGAGGACCATCAGCGCTGGCTGAAGGGCGAAGGGGTCTACTAGGTATGCTGG-TGCC 558
Db 424 TTGAGGACCATCAGCGCTGGCTGAAGGGCGAAGGGGTCTACTAGGTATGCTGGTTGCC 483
Qy 559 TGCCTGTATACAGCCCTGCGCGGGAGTTTCATCAGGGCTCTGGTGTGAGG-CCGCC 617
Db 484 TGCCTGTATACAGCCCTGCGCGGGAGTTTCATCAGGGCTCTGGTGTGAGGCGCCGCC 543
Qy 618 AGGCCCTTTGAGCGGGGCAAGNATACCTCTCAAGTGTGAGGAGACTCAGAG-AGT 676
Db 544 AGGCCCTTTGAGCGGGGCAAGNATACCTCTCAAGTGTGAGGAGACTCAGAGAAAGT 603

Qy 677 GGGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTTAGACCAAGGGGAGCATT-CCAG 735
Db 604 GGGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTTAGACCAAGGGGAGCATTCCAG 663
Qy 736 CATCAGAGATGACACGGATCGCAGGCTGATTGAGAAGAACAGATGAGTGACGGC-AGA 794
Db 664 CATCAGAGATGACACGGATCGCAGGCTGATTGAGAAGAACAGATGAGTGACGGCAAGA 723
Qy 795 AGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTCTCCAGAAAGAGGGGCCGAGA 854
Db 724 AGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTCTCCAGAAAGAGGGGCCGAGA 783
Qy 855 AAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGAGGGG 914
Db 784 AAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGAGGGG 843
Qy 915 ANAGTTAACTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGYNKSGGAAAAAGWG 974
Db 844 AGAGTTAACTGCTGGCTGTGAGTCCCTTGTGGAATATAAGGGGGTGTAGTGGAAACAGTGG 903
Qy 975 TACTAACCCACGATTCTGAGCCCTGAGTATGCTGGACATTGATGCTAAACATGAC 1029
Db 904 TACTAACCCACGATTCTGAGCCCTGAGTATGCTGGACATTGATGCTAAACATGAC 958
RESULT 6
ACN39561
ID ACN39561 standard; cDNA; 963 BP.
XX ACN39561;
AC ACN39561;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA325764, SEQ ID NO:3785.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR P-PSDB; ABM81468.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 3785; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and

| Query Match | 33.4% | Score 403; | DB 2; | Length 464; |
|-----------------------|---|-------------------|------------|------------------|
| Best Local Similarity | 96.2%; | Pred. No. 1e-109; | | |
| Matches 436; | Conservative | 9; | Mismatches | Indels 3; Gaps 3 |
| 752 | GATCGCCAGGCTGATTGACGAACAAAGATGATGTCACGCGC-AGAAGGAGGAGCTCCAGAA | 810 | | |
| 1 | GATCGCCAGGCTGATTGACGAACAAAGATGATGTCACGCGGGAAGAGAGGAGGAGCTCCAGAA | 60 | | |

QY 811 GAGCTTAAACATCTGACCTCTCCAGAGAAGGGGCGGAGAAAGAGAGCTCTAATA 870
DB 61 GAGCTTAAACATCTGACCTCTCCAGAGAAGGGGCGGAGAAAGAGAGCTCTAATA 120
QY 871 AGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAGGGGAGTAACTCTGCTGG 930
DB 121 AGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAGGGGAGTAACTCTGCTGG 180
QY 931 CTGTGANTCCCTTGTGGATATATAAGGGGYSKGGGAAAGHGGTACTAACCCAGATTC 990
DB 181 CTGTGAGTCCCTTGTGGATATATAAGGGGTAAGTGGGAAAGTGGTACTAACCCAGATTC 240
QY 991 TGAGCCCTGAGTATGCTGGAATGATGCTAAACATGACCATGCTTGGGATGCTCTAGC 1050
DB 241 TGAGCCCTGAGTATGCTGGAATGATGCTAAACATGACCATGCTTGGGATGCTCTAGC 300
QY 1051 TGGTCTGGGATAGCTGGAGCACTTACTCAGTGGCTGGTGAATGACACCTCAGAAGGA 1110
DB 301 TGGTCTGGGATAGCTGGAGCACTTACTCAGTGGCTGGTGAATGACACCTCAGAAGGA 360
QY 1111 ATGATGCTATAGAGAGA-GAGAGAGTGTACTGCCAGGTCTTTGACAGATGTAATTC 1169
DB 361 ATNAGTGTATAGAGAGAGAGAGAGTGTACTGCCAGGTCTTTGACAGATGTAATTC 420
QY 1170 TCATTCAATTAATAA-GTTTCAGTGTGTTTGGTTAA 1201
DB 421 TCATTCAATTAAGTTTTCAGTGTGTTTGGTTAA 453

RESULT 13

AAC03762
ID AAC03762 standard; cDNA; 500 BP.

XX AC AAC03762;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein 5' EST, SEQ ID NO: 3760.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.

XX OS Homo sapiens.

XX FN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-00200610.

XX PR 26-FEB-1999; 99US-0122487P.

XX PA (GBST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-500381/45.

XX DR P-ESDB; AAG03756.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX PS Claim 1; SEQ ID NO 3760; 71pp + Sequence Listing; English.

XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. An ORF has been identified within the
XX CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX CC derived from 30 different tissues. EST sequences usually correspond
XX CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX CC well suited for isolating cDNA sequences derived from the 5' ends of
XX CC mRNAs and even in those cases where longer cDNA sequences have been

CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.

SQ Sequence 500 BP; 100 A; 155 C; 133 G; 112 T; 0 U; 0 Other;

Query Match 33.4%; Score 402.2; DB 3; Length 500;
Best Local Similarity 96.8%; Pred. No. 1.9e-109;
Matches 485; Conservative 0; Mismatches 8; Indels 8; Gaps 7;

QY 26 CACGTGACCGCTGACTCGGGGGTCTCCACTATCGCTTACTACTCCCTCTCGAGGAA 85
DB 1 CACGTGACCGCTGACTCGGGGGTCTCCACTATCGCTTACTACTCCCTCTCGAGGAA 60
QY 86 CCGGCGGATATGGCTGCGCGTGTGCCCGCGCGCATTTCTCTCCCGCTGCTT-CCCTT 144
DB 61 CCGGCGGATATGGCTGCGCGTGTGCCCGCGCGCATTTCTCTCCCGCTGCTTCCCTT 120
QY 145 CTCTGGGCTTCTCTGCTCTCTCTCCGCTCCGATGCGCGAGCGGCTGCACACCA-GGC 203
DB 121 CTCTGGGCTTCTCTGCTCTCTCTCCGCTCCGATGCGCGAGCGGCTGCACACCAAGGCG 180
QY 204 GCCCTTCCCTGGATACGGTCACTTTCTACAAGGTCTATCCCAAAGCAAGTTCGTC-TG 262
DB 181 GCCCTTCCCTGGATACGGTCACTTTCTACAAGGTCTATCCCAAAGCAAGTTCGTC-TG 240
QY 263 GTGAAGTTCGACACCCAGTACCCTACGTCGAGAGCAGGATGAGTTCAGCGCTCTT-CT 321
DB 241 GTGAAGTTCGACACCCAGTACCCTACGTCGAGAGCAGGATGAGTTCAGCGCTCTTGTCT 300
QY 322 GAAACCTCGGCTTCCAGCGATGATCTCTTGTGGCAGAGTGGGATCTCAGATTATGTG 381
DB 301 GAAACCTCGGCTTCCAGCGATGATCTCTTGTGGCAGAGTGGGATCTCAGATTATGTG 360
QY 382 ACAAGCTGAACATGGAGCTGAGTGAGAAATCAAGCTGGACAAGAGAGCTACCCA-TCT 440
DB 361 GACAAGCGAACATGGAGCTGAGTGAGAAATCAAGCTGGACAAGAGAGCTACCCAAGTCT 420
QY 441 TCTACCTCTTCCGGGATGGGACTTTGAGACCCAGTCCCATACACTGGGCGAGT--AG 498
DB 421 TCTACCTCTTCCGGGATGGGACTTTGAGAA-CCAGTCCCATACACTGGGCGAGTAAAG 479
QY 499 GTTGGAGCCATCCAGCGCTGG 519
DB 480 GTTGGAGCCATCCAGCGCTGG 500

RESULT 14

AAS83797
ID AAS83797 standard; cDNA; 2087 BP.

XX AC AAS83797;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #19601.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX XX 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG19610.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 1; SEQ ID NO 19601; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 2087 BP; 556 A; 476 C; 549 G; 506 T; 0 U; 0 Other;
XX
XX Query Match 30.6%; Score 369; DB 5; Length 2087;
XX Best Local Similarity 82.8%; Pred. No. 3.4e-99;
XX Matches 893; Conservative 5; Mismatches 107; Indels 74; Gaps 39;
XX
XX 197 CCAAGGGCCCTCCCTGGATACGGTCACATCTTCTACAAAGTCATTTCCCAAAAGCAAGTT 256
XX 619 CAAGGGCCCTTCCCTGGATACGGTCACATCTTCTACAAAGTCATTTCCCAAAAGCAAGTT 678
XX
XX 257 CGTC-TGGTGAAGTTTCGACACCCAGTACCCCTACGGTGAGAAGCAGGATGAGTTCAAGCG 315
XX 679 CGTCTTGTGAAGTTTCGACACCCAGTACCCCTACGGTGAGAAGCAGGATGAGTTCAAGCG 738
XX
XX 316 TCTT-CTGAAACTCGGCTTCAG-CGATGATCTCTT-GTGGGAGAGGTGGGATCTCA 372
XX 739 TTTTGTGAAACTCGGCTTCAGCGGATGATCTCTGGGTGGGAGGTGGGATCTCA 798
XX
XX 373 GATTAT-GTGCAAGCTCAACAT-CGAGCTGAG-TGAGAAATACAACT-GGACAAAGAG 428
XX 799 GATTATGTGCAAGCTCAACATGAGGAGCTGAGTTGAGAAATACAACTGGGACAAAGAG 858
XX
XX 429 AGCTACCCA--TCTTCTACCTTCTCCGGATGGGACTTTT--GAGAAACCCAGTCCCATAC 484
XX 859 AGCTACCCAGTCTTCTTACCTTCTCCGGATGGGACTTTTGGAGAACCCAGTCCCAT 918
XX
XX 485 AC---TGGGCGAGTTA---GGTGGAGCATTCAGCGCTGGCTGAA---GGGCGAAGGG 535
XX 919 ACACCTGGGGGAGTTAAAGGTTGGAGGACATTCAGCGCTGGCTGAAAGGGGGCAAGGG 978
XX
XX 536 TCTACTAGGT--ATGCTGTGGCTGCTGCTGTATACGA-----CGCCCTGGCGCGGG 585
XX 979 TCTACTAGGTGATGCGCTGGAGTTGGTCTGCACTGTATACGTACGCGCTTGGCGGGG 1038
XX
XX 586 AGTTCA--TCAGGGCCTCTGGTGTGGAGGCC--GCCAGGCCCTTTGAAGCAGGGG--CA 639
XX

Db 1039 AATTCAATCAGGGGCTTCTGGTGTGGAGGGCCCGCCAGGCCCTTTTGAAGCAGGGGCACA 1098
Qy 640 AGATAACCTCTCAA--GTGTGAAGGAGACTCAGA---AGAGTGGGGCCGAGCAATACCTGA 694
Db 1099 AGATTAACCTTCAAAGTGTGAAAGGAGACTCAGAAAGAGTGGGGCCGAGCAATACCTGA 1158
Qy 695 AGATCATGGGAAAGATCTTAGA--CCAAAGGGAGCACATT--CCAGCATCAGAGATGACACG 751
Db 1159 AGATCATGGGAAAGATCTTAGACCCAAAGGGAGGACTTCCCAGCATCAGAGATGACAC 1218
Qy 752 GATGCC--CAGGCTGATTGAGAAGAA--CAAGATGAGTGACGGC--AGNAGCAGAGCT-C 805
Db 1219 GGRATCGCTAGGGCTGATTGAGAAGAACCAAGATGAGTGACGGGAGAGAGAGCTCC 1278
Qy 806 CAGAAGAGCTTAA--CATCTGACTGCTTCCAGA---AGAAGGGGCGCCGAGAAAGAGG 860
Db 1279 CAGAAGAGCTTAAACCATCTGACTGCTTCCAGAAAGAGGGGCGCCGAGAAAGAGG 1338
Qy 861 AGCTGTAAA--AAGGCTGCTGTGATTTT--CCAGGGTTTGGTGGGGTAGGGA--GGGGAN 916
Db 1339 AGCTGTAAAACAAGGCTGCTGTGATTTTCCAGGGTTTGGTGGGAGTATGACGGGAGA 1398
Qy 917 AGTTAACTGCTGCTGCTGANTCCCTTGTGAAATATAAGGGGGYMSKGGGAAAGWGG-- 974
Db 1399 GTTAAACCTGCTGCTGAGTCCCTTGTGGAATATAAGGGGGTAGTGGGAAAAGTGGGT 1458
Qy 975 -TACTAACCCACGATTCTGAGCCCTG-AGTATGCTCGAC-ATTGATGCTAACATG--AC 1029
Db 1459 ACTAAACCCCGGATTTGAGCCCTTGAAGTATGCTGGACAAATTGATGCTAACATGAC 1518
Qy 1030 CATGCTTTGGGATGCTCTTAGC--TGGTCTGGGATAGCTGGAGCACTTACTCA--GGTGG 1085
Db 1519 ATGCTTTGGGATGCTCTTAACCTGGGCTGGGGATAGCTGGAGCACTTACTCAAGGTTGG 1578
Qy 1086 CTGTGAAATGACACCTCAGAA--GGATGAGTCTATAGAGGA--GAGAGGAGTGTAC 1142
Db 1579 CTGTGAAATGACACCTCAGAAAGGAATTAAGTGTCTATAGAGGAGAGGAGTGTAC 1638
Qy 1143 TGCCAGGCTTTTGACAGATGTAATTTCTCATTCAATTAAAGTTTCAGTGTTTTGGTTAA 1201
Db 1639 TGCCAGGCTTTTGACAGGTTGTAATTTCTCATTCAATTAAAGTTTCAGTGTTTTGGTTA 1697
XX
XX RESULT 15
XX ID ADT95875 standard; cDNA; 395 BP.
XX AC ADT95875;
XX DT 16-DEC-2004 (first entry)
XX DE Colon cancer associated human cDNA sequence #1394.
XX KW Colon cancer; T cell; tumour protein; C6348; C6358; C6375; C6405; C6368;
XX KW humoral immune response; cellular immune response; cytostatic;
XX OS immunostimulant; human; ss.
XX OS Homo sapiens.
XX FN US2003087818-A1.
XX PD 08-MAY-2003.
XX PF 01-FEB-2002; 2002US-00066543.
XX PR 02-FEB-2001; 2001US-0267400P.
XX PR 07-FEB-2001; 2001US-0267382P.
XX PR 11-MAY-2001; 2001US-0290322P.
XX PR 12-JUL-2001; 2001US-0305265P.
XX PR 16-AUG-2001; 2001US-0313077P.
XX PA (CORI-) CORIXA CORP.
XX

PI Jiang Y, Chenault RA, Xu J, Indrias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX WPI; 2003-040540/03.
DR
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1394; 87pp; English.
XX
XX The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridizes to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 395 BP; 106 A; 88 C; 128 G; 70 T; 0 U; 3 Other;

Query Match 27.3%; Score 328.8; DB 11; Length 395;
Best Local Similarity 97.5%; Pred. No. 1.5e-87;
Matches 385; Conservative 0; Mismatches 5; Indels 5; Gaps 5;

QY 506 CCATCCAGCGCTGGCTGAAGGGGCAAGGGGCTACCTAGGTATGCTGG-TGCCCTGCTG 564
Db 1 CCATCCAGCGCTGGCTGAAGGGGCAAGGGGCTACCTAGGTATGCTGGTGGAGCCCGCAGGCC 60
QY 565 TATACGAGCCCTGGCCGGGAGTTTCATCAGGGCCCTCTGGTGTGAGG-CGCGCCAGGCC 623
Db 61 TATACGAGCCCTGGCCGGGAGTTTCATCAGGGCCCTCTGGTGTGAGGCGCCAGGCC 120
QY 624 TCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAAG-AGTGGGCG 682
Db 121 TCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCANAAGAAGTGGGCG 180
QY 683 AGCATACCTGAGATCATGGGAGATCTTAGACCAAGGGGAGCACTT-CCAGCATCAG 741
Db 181 AGCAATACCTGAGATCATGGGAGATCTTAGACCAAGGGGAGCACTTCCAGCATCAG 240
QY 742 AGATGACACGGATCCCGAGGCTGATTGAGAAGCAAGATGATGAC-CGCAGAGGAGG 800
Db 241 AGATGACACGGATCCCGAGGCTGATTGAGAAGCAAGATGATGATGACCGGAGAGAGG 300
QY 801 AGCTCCAGAAGAGCTTAAACATCTCTGACTGCTTCCAGAAGAGGGGCCGAGAAAGAGG 860
Db 301 AGCTCCANAAGAGCTTAAACATCTCTGACTGCTTCCAGAAGAGGGGCCGANAAGAGG 360
QY 861 AGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTT 895
Db 361 AGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTT 395

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:43:42 ; Search time 190.034 Seconds
(without alignments)
10375.608 Million cell updates/sec

Title: US-08-731-499-7
Perfect score: 1205
Sequence: 1 GCGCGGTGAGTCGCCCCC.....TCAGTGTGTTGGTTAAGTGG 1205

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 1201 | 99.7 | 1205 | 2 | Sequence 7, Appli |
| 2 | 1201 | 99.7 | 1205 | 4 | Sequence 7, Appli |
| 3 | 810.8 | 67.3 | 962 | 4 | Sequence 4780, Ap |
| 4 | 742.8 | 61.6 | 1364 | 4 | Sequence 16522, A |
| 5 | 402.2 | 33.4 | 500 | 4 | Sequence 3760, Ap |
| 6 | 400.4 | 33.2 | 601 | 4 | Sequence 169883, |
| 7 | 191.2 | 15.9 | 363 | 4 | Sequence 8299, Ap |
| 8 | 70.8 | 5.9 | 7218 | 1 | Sequence 14, Appl |
| 9 | 52 | 4.3 | 7218 | 1 | Sequence 14, Appl |
| 10 | 42.2 | 3.5 | 43267 | 4 | Sequence 17117, A |
| 11 | 37.8 | 3.1 | 289 | 3 | Sequence 17, Appl |
| 12 | 37.8 | 3.1 | 289 | 3 | Sequence 17, Appl |
| 13 | 37.8 | 3.1 | 14205 | 4 | Sequence 14464, A |
| 14 | 37.2 | 3.1 | 1526 | 3 | Sequence 63, Appl |
| 15 | 37.2 | 3.1 | 86936 | 4 | Sequence 17314, A |
| 16 | 37 | 3.1 | 27630 | 4 | Sequence 16362, A |
| 17 | 37 | 3.1 | 27684 | 4 | Sequence 12504, A |
| 18 | 36.8 | 3.1 | 53526 | 3 | Sequence 2, Appli |
| 19 | 36.8 | 3.1 | 53577 | 3 | Sequence 1, Appli |
| 20 | 26.6 | 3.0 | 4749 | 4 | Sequence 189, App |
| 21 | 36.4 | 3.0 | 36954 | 4 | Sequence 15304, A |
| 22 | 36.2 | 3.0 | 35688 | 4 | Sequence 16873, A |
| 23 | 36.2 | 3.0 | 68580 | 4 | Sequence 15844, A |
| 24 | 35.6 | 3.0 | 187595 | 4 | Sequence 15446, A |
| 25 | 35.4 | 2.9 | 2810 | 3 | Sequence 6, Appli |
| 26 | 35.4 | 2.9 | 3225 | 1 | Sequence 45, Appl |
| 27 | 35.4 | 2.9 | 3225 | 5 | Sequence 91, Appl |

C 28 35.2 2.9 296 4 US-09-216-393B-97 Sequence 97, Appl
C 29 35.2 2.9 20662 4 US-09-949-016-16717 Sequence 16717, A
C 30 35.2 2.9 43435 4 US-09-949-016-12909 Sequence 12909, A
31 35 2.9 364 4 US-09-621-976-17202 Sequence 17202, A
32 35 2.9 2338 4 US-09-582-337-1 Sequence 1, Appli
33 35 2.9 76767 4 US-09-949-016-12147 Sequence 12147, A
34 35 2.9 76767 4 US-09-949-016-17361 Sequence 17361, A
35 34.8 2.9 375 4 US-09-513-999C-13614 Sequence 13614, A
36 34.8 2.9 9168 4 US-09-687-731-11 Sequence 11, Appl
37 34.4 2.9 289 3 US-09-007-005-17 Sequence 17, Appl
38 34.4 2.9 289 3 US-09-244-796-17 Sequence 17, Appl
39 34.4 2.9 601 4 US-09-949-016-51960 Sequence 51960, A
40 34.4 2.9 12368 4 US-09-949-016-13243 Sequence 13243, A
41 34.4 2.9 247299 4 US-09-949-016-17590 Sequence 17590, A
C 42 34.2 2.8 430 4 US-09-621-976-16656 Sequence 16656, A
43 34.2 2.8 62386 4 US-09-949-016-12823 Sequence 12823, A
44 34.2 2.8 390890 4 US-09-949-016-14720 Sequence 14720, A
C 45 34 2.8 39085 4 US-09-949-016-14479 Sequence 14479, A

ALIGNMENTS

RESULT 1
US-08-680-395-7
; Sequence 7, Application US/08680395
; Patent No. 5892010
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W.
; APPLICANT: Collins, Colin
; APPLICANT: Hwang, Soo-in
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowbel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: Genes from the 20ql3 Amplicon and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,395
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-0689000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: 1..1205
; LOCATION: /note= "cdna clone lb4 for a serine
; OTHER INFORMATION: threonine kinase"
; OTHER INFORMATION:
US-08-680-395-7

| | | | | | | | | | |
|---|-----|---|------|--|--|--|--|--|--|
| Query Match | | | | | | | | | |
| Best Local Similarity 99.7%; Score 1201; DB 2; Length 1205; | | | | | | | | | |
| Matches 1205; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| Qy | 1 | GCGCGGTGAGTCGCGCCCCCAGCTACAGTACCGCTGACCTCGGGGGGTTCTCCACATATC | 60 | | | | | | |
| Db | 1 | GCGGCGGTGAGTCGCGCCCCCAGCTACAGTACCGCTGACCTCGGGGGGTTCTCCACATATC | 60 | | | | | | |
| Qy | 61 | GCTTACCTACCTCCCTCTGCAGGAACCGCGGATATGCTGCCGCTGTCGCCCGCGCGC | 120 | | | | | | |
| Db | 61 | GCTTACCTACCTCCCTCTGCAGGAACCGCGGATATGCTGCCGCTGTCGCCCGCGCGC | 120 | | | | | | |
| Qy | 121 | ATTTCTCTCCCGCTGCTTCCCTTCTCCTGGGCTTCTGCTCTCCCTCCGCTCCGATGGC | 180 | | | | | | |
| Db | 121 | ATTTCTCTCCCGCTGCTTCCCTTCTCCTGGGCTTCTGCTCTCCCTCCGCTCCGATGGC | 180 | | | | | | |
| Qy | 181 | GGCAGCGCTGACACCAAGGCGCTTCCCTTGGATACGGTCACTTTCTACAAGGTCA | 240 | | | | | | |
| Db | 181 | GGCAGCGCTGACACCAAGGCGCTTCCCTTGGATACGGTCACTTTCTACAAGGTCA | 240 | | | | | | |
| Qy | 241 | TTCCCAAAGCAAGTTCGTTGGTGAAGTTCCGACCCAGTACCCCTACGGTGAGAGCA | 300 | | | | | | |
| Db | 241 | TTCCCAAAGCAAGTTCGTTGGTGAAGTTCCGACCCAGTACCCCTACGGTGAGAGCA | 300 | | | | | | |
| Qy | 301 | GGATGAGTTCAAGCGTCTTCTGAAACTTCGGCTTCCAGCGATGATCTTGGTGGCAGAG | 360 | | | | | | |
| Db | 301 | GGATGAGTTCAAGCGTCTTCTGAAACTTCGGCTTCCAGCGATGATCTTGGTGGCAGAG | 360 | | | | | | |
| Qy | 361 | GTGGGGTCTCAGATTATGTGCAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGG | 420 | | | | | | |
| Db | 361 | GTGGGGTCTCAGATTATGTGCAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGG | 420 | | | | | | |
| Qy | 421 | ACAAAGAGAGTACCCATCTTCTACCTTCCGGATGGGACTTTGAGACCCAGTCCC | 480 | | | | | | |
| Db | 421 | ACAAAGAGAGTACCCATCTTCTACCTTCCGGATGGGACTTTGAGACCCAGTCCC | 480 | | | | | | |
| Qy | 481 | ATACACTGGGCGAGTTAGTTGGAGCCATCCAGCGCTGGCTGAAAGGGGCAAGGGTCTAC | 540 | | | | | | |
| Db | 481 | ATACACTGGGCGAGTTAGTTGGAGCCATCCAGCGCTGGCTGAAAGGGGCAAGGGTCTAC | 540 | | | | | | |
| Qy | 541 | CTAGGTATGCTGTGCTGCTGTATACGACGCCCTGGCCGGGGAGTTATCAGGGCCT | 600 | | | | | | |
| Db | 541 | CTAGGTATGCTGTGCTGCTGTATACGACGCCCTGGCCGGGGAGTTATCAGGGCCT | 600 | | | | | | |
| Qy | 601 | CTGTGTGAGGCGCCAGGCGCTTCTGAAGCGGGCAAGATAACCTCTCAAGTGTGAA | 660 | | | | | | |
| Db | 601 | CTGTGTGAGGCGCCAGGCGCTTCTGAAGCGGGCAAGATAACCTCTCAAGTGTGAA | 660 | | | | | | |
| Qy | 661 | GGAGACTCAGAAAGTGGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAA | 720 | | | | | | |
| Db | 661 | GGAGACTCAGAAAGTGGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAA | 720 | | | | | | |
| Qy | 721 | GGGAGACATCTTCAGCATCAGAGATGACACCGATCGCCAGCTGATTGAGAAACAGA | 780 | | | | | | |
| Db | 721 | GGGAGACATCTTCAGCATCAGAGATGACACCGATCGCCAGCTGATTGAGAAACAGA | 780 | | | | | | |
| Qy | 781 | TGAGTGAGGAGAGGAGGCTCCAGAGCTTAAACATCTGACTGCTCCAGAA | 840 | | | | | | |
| Db | 781 | TGAGTGAGGAGAGGAGGCTCCAGAGCTTAAACATCTGACTGCTCCAGAA | 840 | | | | | | |
| Qy | 841 | GAAAGGGGCCAGAAAGAGGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGTTTGGTG | 900 | | | | | | |
| Db | 841 | GAAAGGGGCCAGAAAGAGGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGTTTGGTG | 900 | | | | | | |
| Qy | 901 | GGGTAGGAGGGGAGTAACTCTGCTGGCTGTGANTCCCTTGTGGAAATATAAGGGGY | 960 | | | | | | |
| Db | 901 | GGGTAGGAGGGGAGTAACTCTGCTGGCTGTGANTCCCTTGTGGAAATATAAGGGGY | 960 | | | | | | |
| Qy | 961 | MSKGGGAAAGWGGTAAACCCAGATTCTGAGCCCTGAGTATGCCCTGGACATTGATGC | 1020 | | | | | | |
| Db | 961 | MSKGGGAAAGWGGTAAACCCAGATTCTGAGCCCTGAGTATGCCCTGGACATTGATGC | 1020 | | | | | | |

RESULT 2
US-08-892-695-7
; Sequence 7, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20013 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1205
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:1b4
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (916)
; OTHER INFORMATION: n is A, C, G, T, or U
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (937)
; OTHER INFORMATION: n is A, C, G, T, or U
US-08-892-695-7

| | | | | | | | | | |
|---|-----|---|-----|--|--|--|--|--|--|
| Query Match | | | | | | | | | |
| Best Local Similarity 99.7%; Score 1201; DB 4; Length 1205; | | | | | | | | | |
| Matches 1205; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| Qy | 1 | GCGCGGTGAGTCGCGCCCCCAGTACAGTACCGCTGACCTCGGGGGTTCCTCCACTATC | 60 | | | | | | |
| Db | 1 | GCGCGGTGAGTCGCGCCCCCAGTACAGTACCGCTGACCTCGGGGGTTCCTCCACTATC | 60 | | | | | | |
| Qy | 61 | GCTTACCTACCTCCCTCTGCAGNAACCGGGGATATGGCTGCCCGCGCGCGC | 120 | | | | | | |
| Db | 61 | GCTTACCTACCTCCCTCTGCAGNAACCGGGGATATGGCTGCCCGCGCGCGC | 120 | | | | | | |
| Qy | 121 | ATTTCTCTCCCGCTGCTTCCCTTCTCCTGGGCTTCTGCTCTCCCTCCGCTCCGCATGGC | 180 | | | | | | |
| Db | 121 | ATTTCTCTCCCGCTGCTTCCCTTCTCCTGGGCTTCTGCTCTCCCTCCGCTCCGCATGGC | 180 | | | | | | |
| Qy | 181 | GGCAGCGGCTGACACCAAGGCGCCCTTCCCTCGGATACGGTCACTTTCTACAAGGTCA | 240 | | | | | | |


```
Db 181 GCACGGGCTGCACACCAAGCGCCCTTCCCTGGATACGGTCACTTCTACAGGTCA 240
QY 241 TTCCCAAAAGCAAGTTGCTGCTGGTGAAGTTGACACCCAGTACCCCTACGGTGAAGA 300
Db 241 TTCCCAAAAGCAAGTTGCTGCTGGTGAAGTTGACACCCAGTACCCCTACGGTGAAGA 300
QY 301 GGATGAGTTCAAGCGTCTTCTGAAAACCTCGGCTCCAGCGATGATCTCTTGGTGCAG 360
Db 301 GGATGAGTTCAAGCGTCTTCTGAAAACCTCGGCTCCAGCGATGATCTCTTGGTGCAG 360
QY 361 GTGGGGATCTCAGATTATGTGACAGCTGAAACATGAGCTGAGTGAAGAAATACAGCTGG 420
Db 361 GTGGGGATCTCAGATTATGTGACAGCTGAAACATGAGCTGAGTGAAGAAATACAGCTGG 420
QY 421 ACAAGAGAGCTACCCATCTTCTACTCTTCCGGATGGGACTTTGAGAACCCAGTCCC 480
Db 421 ACAAGAGAGCTACCCATCTTCTACTCTTCCGGATGGGACTTTGAGAACCCAGTCCC 480
QY 481 ATACACTGGGCGAGTTAGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGTCTAC 540
Db 481 ATACACTGGGCGAGTTAGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGTCTAC 540
QY 541 CTAGTATGCTGCTGCTGCTGCTATACAGCGCTGCGCGGGAGTTCTACAGGGCT 600
Db 541 CTAGTATGCTGCTGCTGCTGCTATACAGCGCTGCGCGGGAGTTCTACAGGGCT 600
QY 601 CTGCTGTGGAGCGCCAGCGCTCTTGAAGCAGGGCAAGATAACCTCTCAAGTGTGAA 660
Db 601 CTGCTGTGGAGCGCCAGCGCTCTTGAAGCAGGGCAAGATAACCTCTCAAGTGTGAA 660
QY 661 GGAGACTCAGAAGAGTGGCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACAA 720
Db 661 GGAGACTCAGAAGAGTGGCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACAA 720
QY 721 GGGGAGCACTTCCAGCATCAGAGATGACACGATCGCAGGCTGATGAGAAGAACAGA 780
Db 721 GGGGAGCACTTCCAGCATCAGAGATGACACGATCGCAGGCTGATGAGAAGAACAGA 780
QY 781 TGAGTGACGGCAAGAGGAGCTCCAGAAGCTTAAACATCTCTGACTGCTTCCAGAA 840
Db 781 TGAGTGACGGCAAGAGGAGCTCCAGAAGCTTAAACATCTCTGACTGCTTCCAGAA 840
QY 841 GAAGGGGCGGAGAAAGAGGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTTGGTG 900
Db 841 GAAGGGGCGGAGAAAGAGGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTTGGTG 900
QY 901 GGGGTAGGGAGGGANAGTTAACCTGCTGGCTGANTCCCTTGTGGGAATATAAGGGGY 960
Db 901 GGGGTAGGGAGGGANAGTTAACCTGCTGGCTGANTCCCTTGTGGGAATATAAGGGGY 960
QY 961 MSKGGGAAAAGWGGTACTAACCCAGATTCTGAGCCCTGAGTATGCTTGGACATTGATGC 1020
Db 961 MSKGGGAAAAGWGGTACTAACCCAGATTCTGAGCCCTGAGTATGCTTGGACATTGATGC 1020
QY 1021 TAAATGACCATGCTTGGGATGTCTAGCTGGTCTGGGGATGAGCTGGAGCACTTACTCA 1080
Db 1021 TAAATGACCATGCTTGGGATGTCTAGCTGGTCTGGGGATGAGCTGGAGCACTTACTCA 1080
QY 1081 GGTGCTGCTGAAATGACACCTCAGAAGAAATGAGTGTATAGAGAGAGAGAGTGT 1140
Db 1081 GGTGCTGCTGAAATGACACCTCAGAAGAAATGAGTGTATAGAGAGAGAGAGTGT 1140
QY 1141 ACTGCCCAAGGCTTTTGACAGATGTAATTTCTCATTTAAAGTTTCAAGTGTGGTTA 1200
Db 1141 ACTGCCCAAGGCTTTTGACAGATGTAATTTCTCATTTAAAGTTTCAAGTGTGGTTA 1200
QY 1201 AGTGG 1205
Db 1201 AGTGG 1205
```

RESULT 3

US-09-949-016-4780

; Sequence 4780, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR FILING DATE: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4780

; LENGTH: 962

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-4780

Query Match 57.3%; Score 810.8; DB 4; Length 962;

Best Local Similarity 97.8%; Pred. No. 1.9e-233;

Matches 941; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

```
QY 86 CCCGCGCATATGGCTGCGCTGTCGCCCGCCGCGCATTTCTCTCCCGCTGCTT-CCCTT 144
Db 1 CCCGCGCATATGGCTGCGCTGTCGCCCGCCGCGCATTTCTCTCCCGCTGCTTCCCTT 60
QY 145 CTCCTGGGCTTCTGCTCTCTCTCCGCTCCGATGGCGCAGCGGCTGCACACCAA-GGC 203
Db 61 CTCCTGGGCTTCTGCTCTCTCTCCGCTCCGATGGCGCAGCGGCTGCACACCAAGGCG 120
QY 204 GCCCTTCCCTGGATACGGTCACTTCTACAGGTCAATTCCTCCCAAGCAAGTCTGTC-TG 262
Db 121 GCCCTTCCCTGGATACGGTCACTTCTACAGGTCAATTCCTCCCAAGCAAGTCTGTC-TG 180
QY 263 GTGAAGTTCCGACACCCAGTACCCCTAGCGTGAGAAGCAGGATGAGTTCAGCGTCTT-CT 321
Db 181 GTGAAGTTCCGACACCCAGTACCCCTAGCGTGAGAAGCAGGATGAGTTCAGCGTCTTCT 240
QY 322 GAAACTCGGCTTCCAGCGATGATCTCTTGTGGCAGAGTGGGATCTCAGATTAT-GT 380
Db 241 GAAACTCGGCTTCCAGCGATGATCTCTTGTGGCAGAGTGGGATCTCAGATTATGCT 300
QY 381 GACAAGCTGAACATGGAGCTGAGTGAGAAATACAACTGGGCAAAAGAGAGCTACCCA-TC 439
Db 301 GACAAGCTGAACATGGAGCTGAGTGAGAAATACAACTGGGCAAAAGAGAGCTACCCAAGTC 360
QY 440 TTCTACCTCTTCCGGGATGGGACTTTGAGAAACCCAGTCCCATACACCTGGGGCAGTT-AG 498
Db 361 TTCTACCTCTTCCGGGATGGGACTTTGAGAAACCCAGTCCCATACACCTGGGGCAGTTAAG 420
QY 499 GTTGAGCCATCCAGCGCTGCTCAAGGGCAAGGGGTCTACCTAGGTATGCTGG-TGC 557
Db 421 GTTGAGCCATCCAGCGCTGCTCAAGGGCAAGGGGTCTACCTAGGTATGCTGGTTGC 480
QY 558 CTGCTGTATACGACGCGCTGGCGGGGAGTTTCATCAGGGCTCTGGTGTGGAGG-CCGC 616
Db 481 CTGCTGTATACGACGCGCTGGCGGGGAGTTTCATCAGGGCTCTGGTGTGGAGGCGCCGC 540
QY 617 CAGGCCCTTTGAAAGCAGGGGCAAGATAAAGTCTCAAGTGTGAAGGAGACTCAGAAG-AG 675
Db 541 CAGGCCCTTTGAAAGCAGGGGCAAGATAAAGTCTCAAGTGTGAAGGAGACTCAGAAGAG 600
QY 676 TGGGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CCA 734
Db 601 TGGGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTTCCA 660
QY 735 GCATCAGAGATGACACCGATCGCCAGGCTGATTGAGAAAGCAAGATGAGTGAC-GGCAG 793
```

Db 661 GCATCAGAGATGACACGGATCCCGAGTCTGATGAGAAACAAGATGAGTGCAGGGAAG 720
Qy 794 AAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAAAGAGGAGGCGAG 853
Db 721 AAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAAAGAGGAGGCGAG 780
Qy 854 AAGAGGAGCTGTAAAGAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAGGG 913
Db 781 AAGAGGAGCTGTAAAGAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAGGG 840
Qy 914 GANAGTTAACTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGAGGAGGAGGAGG 973
Db 841 GAGAGTTAACTGCTGGCTGTGAGTCCCTTGTGGAATATAAGGGGAGGAGGAGGAGG 900
Qy 974 GTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGAGCATTTGATGCTAAACATGACCATG 1033
Db 901 GTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGAGCATTTGATGCTAAACATGACCATG 960
Qy 1034 CT 1035
Db 961 CT 962

RESULT 4

US-09-949-016-16522
; Sequence 16522, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16522
; LENGTH: 13364
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)..(13364)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16522

Query Match 61.6%; Score 742.8; DB 4; Length 13364;
Best Local Similarity 98.1%; Pred. No. 2.6e-212;
Matches 818; Conservative 5; Mismatches 4; Indels 7; Gaps 7;
Qy 379 GTGCAAGCTGAACATGGAGCTGAGTGAAGATACAAAGCTGGACAAAGAGAGTACCCA- 437
Db 10700 GTGCAAGCTGAACATGGAGCTGAGTGAAGATACAAAGCTGGACAAAGAGAGTACCCA 10759
Qy 438 TCTTCTACCTCTTCGGGATGGGACTTTTGAACCCAGTCCCATACACTGGGGCAGTT- 496
Db 10760 TCTTCTACCTCTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGTTA 10819
Qy 497 AGGTTGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACCTAGTATGCTGG-T 555
Db 10820 AGGTTGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACCTAGTATGCTGGTT 10879
Qy 556 GCCTCCCTGTATACGACGCCCTGGCCGGGAGTTTCATCAGGCGCTCTGTGTGGAGG-CC 614
Db 10880 GCCTCCCTGTATACGACGCCCTGGCCGGGAGTTTCATCAGGCGCTCTGTGTGGAGGCC 10939

Qy 615 GCCAGGCCCTCTTGAAGCAGGGGCAAGATAACTCTCAAGTGTGAAGAGACTCAGAAG- 673
Db 10940 GCCAGGCCCTCTTGAAGCAGGGGCAAGATAACTCTCAAGTGTGAAGAGACTCAGAAGA 10999
Qy 674 AGTGGGCGGAGCATACCTGAAGATCATGGGAAGATCTTAGACCAAGGGGAGCATT-C 732
Db 11000 AGTGGGCGGAGCAATCCTGAAGATCATGGGAAGATCTTAGACCAAGGGGAGCATTCC 11059
Qy 733 CAGCATCAGAGATGACACGGATCCAGAGCTGATTTGAGAAGAAACAAGATGAGTGAC-GGC 791
Db 11060 CAGCATCAGAGATGACACGGATCCAGAGCTGATTTGAGAAGAAACAAGATGAGTGACGGGA 11119
Qy 792 AGAAGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTCCAGAAAGAGGGGCGC 851
Db 11120 AGAAGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTCCAGAAAGAGGGGCGC 11179
Qy 852 AGAAGAGGAGCTTAAAGAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAG 911
Db 11180 AGAAGAGGAGCTTAAAGAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAG 11239
Qy 912 GGGANAGTTAACTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGAGGAGGAG 971
Db 11240 GGGAGAGTTAACTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGAGGAGGAG 11299
Qy 972 WGTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGAGCAATTGATGCTAACATGACCA 1031
Db 11300 TGGTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGAGCAATTGATGCTAACATGACCA 11359
Qy 1032 TGCTTGGGATGCTCTAGCTGGCTGGGAGTAGCTGGAGCACTTACTCAGCTGGCTGGTG 1091
Db 11360 TGCTTGGGATGCTCTAGCTGGCTGGGAGTAGCTGGAGCACTTACTCAGCTGGCTGGTG 11419
Qy 1092 AAATGACACCTCAGAAAGAAATGAGTGTCTATAGAGAGAGAGAGGAGTGTACTGCCAGGT 1151
Db 11420 AAATGACACCTCAGAAAGAAATGAGTGTCTATAGAGAGAGAGAGGAGTGTACTGCCAGGT 11479
Qy 1152 CTTTGACAGATGTAATTTCTCAATCAATTAAGTTTCAAGTTTGGTTAAAGTG 1205
Db 11480 CTTTGACAGATGTAATTTCTCAATCAATTAAGTTTCAAGTTTGGTTAAAGTG 11533

RESULT 5

US-09-513-999C-3760
; Sequence 3760, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 3760
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 70..378
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 70..129
; OTHER INFORMATION: score 6.9
; OTHER INFORMATION: seq AAFSLPILLPLLLIG/FL
US-09-513-999C-3760

Query Match 33.4%; Score 402.2; DB 4; Length 500;

Best Local Similarity 96.8%; Pred. No. 1.2e-110; Matches 485; Conservative 0; Mismatches 8; Indels 8; Gaps 7;

QY 26 CACGTGACCGTGAACCGCGGCTTCCACTATCGCTTACCTACCTCCCTCTGCAGAA 85
DB 1 CACGTGACCGTGAACCGCGGCTTCCACTATCGCTTACCTACCTCCCTCTGCAGAA 60
QY 86 CCCGCGGATATGGCTGCGCTGTCGCCCGCGCGCAATTCCTCCCGCTGCTT-CCCTT 144
DB 61 CCCGCGGATATGGCTGCGCTGTCGCCCGCGCGCAATTCCTCCCGCTGCTTCCCTT 120
QY 145 CTCCTGGGCTTCTCTCTCTCTCCGCTCCGATGCGCGCGCTGCACACAA-GGC 203
DB 121 CTCCTGGGCTTCTCTCTCTCTCCGCTCCGATGCGCGCGCTGCACACAAAGGC 180
QY 204 GCCCTTCCCTGGATACGGTCACTTTCTACAAGGTCAATCCAAAGCAAGTTGTC-TG 262
DB 181 GCCCTTCCCTGGATACGGTCACTTTCTACAAGGTCAATCCAAAGCAAGTTGTC-TG 240
QY 263 GTGAAGTTTCGACACCCAGTACCCCTACCGTGAAGCAGGATGATTCAGGGTCTT-CT 321
DB 241 GTGAAGTTTCGACACCCAGTACCCCTACCGTGAAGCAGGATGATTCAGGGTCTTCT 300
QY 322 GAAACTCGGCTTCCAGCGATGATCTTTGGTGGCAGAGTGGGATCTCAGATTATGTG 381
DB 301 GAAACTCGGCTTCCAGCGATGATCTTTGGTGGCAGAGTGGGATCTCAGATTATGTG 360
QY 382 ACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGCTACCCA-TCT 440
DB 361 GACAGCGAATCATGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGCTACCCAGTCT 420
QY 441 TCTACTCTTCCGGATGCGGACTTTGAGAACCCAGTCCCATACACTGGGCGAGTT--AG 498
DB 421 TCTACTCTTCCGGATGCGGACTTTGAGAA-CCAGTCCCATACACTGGGCGAGTTAAAG 479
QY 499 GTTGAGGCATCCAGCGCTGG 519
DB 480 GTTGAGGCATCCAGCGCTGG 500

RESULT 6
US-09-949-016-169883
; Sequence 169883, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169883
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-169883

Query Match 33.2%; Score 400.4; DB 4; Length 601;
Best Local Similarity 98.0%; Pred. No. 4.6e-110; Matches 478; Conservative 1; Mismatches 1; Indels 7; Gaps 7;

QY 379 GTGACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGCTACCCA- 437
DB 114 GTGACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGCTACCCA 173

QY 438 TCTTCTACCTCTTCGGGATGGGACTTTTGAGAACCCAGTCCCATACACTGGGCGAGTT- 496
DB 174 TCTTCTACCTCTTCGGGATGGGACTTTTGAGAACCCAGTCCCATACACTGGGCGAGTTA 233
QY 497 AGGTTGGAGCCATCCAGCGCTGGCTGGAAGGGGCAAGGGGTCTACCTAGGTATGCTGG-T 555
DB 234 AGGTTGGAGCCATCCAGCGCTGGCTGGAAGGGGCAAGGGGTCTACCTAGGTATGCTGGTT 293
QY 556 GCCTGCTGTATACGACGCCCTGGCGGGGAGTTTCATCAGGGCTCTCTGGTGTGAGG-CC 614
DB 294 GCCTGCTGTATACGACGCCCTGGCGGGGAGTTTCATCAGGGCTCTCTGGTGTGAGGCCC 353
QY 615 GCCAGGCCCTCTTCAAGCAGGCGCAAGATACCTCTCAAGTGTGAAGGAGACTCAGAAG- 673
DB 354 GCCAGGCCCTCTTGAAGCAGGCGCAAGATACCTCTCAAGTGTGAAGGAGACTCAGAAGA 413
QY 674 AGTGGGCGGACCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-C 732
DB 414 AGTGGGCGGACCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTTCC 473
QY 733 CAGCATCAGATGACACGGATCCCGAGTTCGAGTGTGAGAAACAAGATGATGAC-GGC 791
DB 474 CAGCATCAGATGACACGGATCCCGAGTTCGAGTGTGAGAAACAAGATGATGACGCGGA 533
QY 792 AGAAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAGGGGGCGG 851
DB 534 AGAAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAGGGGGCGG 593
QY 852 AGAAGAGAG 859
DB 594 AGAAGAGAG 601

RESULT 7
US-09-513-999C-8299
; Sequence 8299, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36881
; SOFTWARE: Patent.pm
; SEQ ID NO 8299
; LENGTH: 363
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 175
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 176
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 188
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 214
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 215

; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-8299

Query Match 15.9%; Score 191.2; DB 4; Length 363;
Best Local Similarity 94.3%; Pred. No. 3.1e-47; Indels 2; Gaps 2;
Matches 215; Conservative 3; Mismatches 8;
Qy 26 CACGTGACCGCTGACTCGGGGGTTCCTCCACTATCGCTTACCTACCTCCCTCTCGAGGAA 85
Db 1 CACGTGACCGCTGACTCGGGGGTTCCTCCACTATCGCTTACCTACCTCCCTCTCGAGGAA 60
Qy 86 CCGGGGATATGGCTGCGCTGTGCGCGCGCGCGGATTTCTCTCCGCTGCTT-CCCTT 144
Db 61 CCGGGGATATGGCTGCGCTGTGCGCGCGCGCGGATTTCTCTCCGCTGCTTCCCTT 120
Qy 145 CTCCTGGGCTTCCTGCTCTCTCGCTCGGATGCGCGCGCTGCACACAA-GGC 203
Db 121 CTCCTGGGCTTCCTGCTCTCTCGCTCGGATGCGCGCGCTGCACACRRGGC 180
Qy 204 GCCCTTCCTCGGATACGGTCACTTTCTACAAGGTCAATTCCTCCAAAGC 251
Db 181 GCCCTTCCTCGGATACGGTCACTTTCTACAAGRATATGGTGACAAGC 228

RESULT 8

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 5.9%; Score 70.8; DB 1; Length 7218;
Best Local Similarity 3.2%; Pred. No. 3e-10;
Matches 12; Conservative 228; Mismatches 132; Indels 0; Gaps 0;
Qy 628 GAACGAGGGCAAGATACCTCTCAAGTGTCAAGGAGACTCAGAAGAGTGGCCGAGCAA 687
Db 1417 RRR 1358
Qy 688 TACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCAGCTTCCAGCATCAGAGATGA 747
Db 1357 RRR 1298
Qy 748 CACGATCGCAGCTGATTTAGAGAAACAAGATGATGACGCGAGAGAGGAGCTCCA 807
Db 1297 RRR 1238
Qy 808 GAAGAGCTTAAACATCTGACTGCTTCCAGAAAGGGGCGCCAGAAAGAGGAGCTCTA 867
Db 1237 RRR 1178
Qy 868 AAAAGGCTGCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGGANAGTTAACCTGC 927
Db 1177 RRR 1118
Qy 928 TGGCTGTGANTCCCTTGTGGAATATAAGGGGYSKGGGAAAAGWGTTACTAACCCAGA 987
Db 1117 RRR 1058
Qy 988 TTCTGAGCCCTG 999
Db 1057 TCCTCGACCTG 1046

RESULT 9

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:

RESULT 12
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihue
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS

```
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match      3.1%; Score 37.8; DB 3; Length 289;
Best Local Similarity 5.2%; Pred. No. 0.39;
Matches 12; Conservative 103; Mismatches 114; Indels 0; Gaps 0;

Qy  2 CGCGGTGAGTCCGCCCCCAGTACAGTACGCGTACGCGGTCTCCACTATCG 61
Db  247 YGYAYGYTYAYCYGYCYAYGYCYGYCYGYCYGYCYGYCYGYCYGYCY 188
Qy  62 CTACTACTACCTCCCTCTCAGAGAACCCGCGGATGCGTGGCGTGTCCCGCGCGCA 121
Db  187 YNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNY 128
Qy  122 TTTCTCTCCCGCTGCTCCCTTCCTGCGGTCTCCTGCTCTCCTGCTCGGATGCG 181
Db  127 YNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNY 68
Qy  182 GCAGCGGCTGCACACCAAGCGCGCTCCCTCCCTGGATACGTCATTTC 230
Db  67 YNYNYNYNYCYAVTYTYGYTYAYTYTYGYTYAYTYGYTYAYTYGYTY 19

RESULT 13
US-09-949-016-14464
; Sequence 14464, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14464
; LENGTH: 14205
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14464

Query Match      3.1%; Score 37.8; DB 4; Length 14205;
Best Local Similarity 52.1%; Pred. No. 3.8;
Matches 111; Conservative 0; Mismatches 97; Indels 5; Gaps 1;

; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; EARLIER FILING DATE: 1998-01-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match      3.1%; Score 37.2; DB 3; Length 1526;
Best Local Similarity 52.6%; Pred. No. 1.6;
Matches 81; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy  73 CCCTCTGAGGAACCCGCGGATATGGCTGCGGTGTCGCCCGCGCGCATTTCTCTCCC 132
Db  31 CTCCCTGCGCTGTGCCCGCGGTTCGCCGGGCGCTGACCCCGACCTCTCCAGCCTGTCCCC 90
Qy  133 GCTGCTTCCCTTCTCCTGGGCTTCTGCTCTCTCCGCTCCGCTGCGGCGGCGGCGCTG 192
Db  91 GTGCGCTCTCTGGGCGCTGTGCCCCACCTCTCCCGGCTGTCCCCACCTCTCCCGGCTT 150
Qy  193 CACACCAAGGCGGCGCTTCCCTCTGGATACGGTCA 226
Db  151 GTCCCCCACCTCCCGGCGCTGTCCCCCTGTCCCC 184

RESULT 15
US-09-949-016-17314
; Sequence 17314, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17314
; LENGTH: 86936
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(86936)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17314

Query Match      3.1%; Score 37.2; DB 4; Length 86936;
Best Local Similarity 51.2%; Pred. No. 17;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 53 CCACATCGCTTACCTACCTCCCTGAGGAAACCGGGGATATGGCTGCGGCTGTGCC 112
Db 95 CAACAATTGCCCTGGTCTCCCCCAAGTGACGGTTCTCGCCCGCCCGCTCCGCT 154
Qy 113 CGCGCGCATTTCTCTCCCGGCTGTTCCCTTCTCTGGGTTCTGCTCTCTCCGCTC 172
Db 155 CGCCCGGCGCTTCCCGGCTCTCCCGGCTCTCCCGGCTCTCCCGGCTCTCCCGCT 214
Qy 173 CGCATGGGGGAGCGGCTTGACACAAAGGCGCCCTTCCCTGGATACGG 222
Db 215 CTCCCGGCGCTTCCCGGCTTCCCGGCTTCCCGGCTTCAACCTAGTGAGGG 264
```

Search completed: July 3, 2005, 16:59:32
Job time : 193.034 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 16:40:17 ; Search time 703.092 Seconds

(without alignments)
10749.639 Million cell updates/sec

Title: US-08-731-499-7

Perfect score: 1205

Sequence: 1 GCGCGGAGTCCGCCCC.....TCAGTGTTCGTTAAGTGG 1205

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|---------------------|
| 1 | 1201 | 99.7 | 1205 | 8 | US-08-731-499-7 |
| 2 | 1062.8 | 88.2 | 1334 | 9 | US-09-925-301-566 |
| 3 | 803.8 | 66.7 | 963 | 15 | US-10-205-823-43 |
| 4 | 803.8 | 66.7 | 963 | 21 | US-10-956-157-1670 |
| 5 | 517.2 | 42.9 | 4529 | 17 | US-10-191-803-273 |
| 6 | 517.2 | 42.9 | 4529 | 18 | US-10-152-319A-2052 |
| 7 | 515.8 | 42.8 | 600 | 21 | US-10-956-157-6905 |

| | | | | | | |
|----|-------|------|--------|----|----------------------|--------------------|
| 8 | 496.6 | 41.2 | 1139 | 17 | US-10-388-934-523 | Sequence 523, App |
| 9 | 404.8 | 33.6 | 438 | 17 | US-10-242-535A-37352 | Sequence 37352, A |
| 10 | 404.8 | 33.6 | 438 | 18 | US-10-085-783A-37352 | Sequence 37352, A |
| 11 | 328.8 | 27.3 | 395 | 14 | US-10-066-543-1394 | Sequence 1394, Ap |
| 12 | 321 | 26.6 | 405 | 10 | US-09-918-995-5850 | Sequence 5850, Ap |
| 13 | 279.2 | 23.2 | 450 | 9 | US-09-864-761-2375 | Sequence 2375, Ap |
| 14 | 271.8 | 22.6 | 345 | 10 | US-09-918-995-18955 | Sequence 18955, A |
| 15 | 252.4 | 20.9 | 255 | 17 | US-10-242-535A-31517 | Sequence 31517, A |
| 16 | 252.4 | 20.9 | 255 | 18 | US-10-085-783A-31517 | Sequence 31517, A |
| 17 | 252 | 20.9 | 255 | 17 | US-10-242-535A-23837 | Sequence 23837, A |
| 18 | 252 | 20.9 | 255 | 18 | US-10-085-783A-23837 | Sequence 23837, A |
| 19 | 225.4 | 18.7 | 281 | 18 | US-10-424-599-16936 | Sequence 16936, A |
| 20 | 223.8 | 18.6 | 248 | 20 | US-10-425-115-120840 | Sequence 120840, A |
| 21 | 220.2 | 18.3 | 360 | 21 | US-10-467-606-3 | Sequence 3, Appli |
| 22 | 218.8 | 18.2 | 270 | 20 | US-10-425-115-59573 | Sequence 59573, A |
| 23 | 213.8 | 17.7 | 383 | 9 | US-09-960-352-14767 | Sequence 14767, A |
| 24 | 186.8 | 15.5 | 197 | 17 | US-10-242-535A-19904 | Sequence 19904, A |
| 25 | 186.8 | 15.5 | 197 | 18 | US-10-085-783A-19904 | Sequence 19904, A |
| 26 | 165 | 13.7 | 165 | 14 | US-10-066-543-2866 | Sequence 2866, Ap |
| 27 | 151.4 | 12.6 | 415 | 9 | US-09-960-352-12723 | Sequence 12723, A |
| 28 | 140 | 11.6 | 210 | 17 | US-10-242-535A-4438 | Sequence 4438, Ap |
| 29 | 140 | 11.6 | 210 | 18 | US-10-085-783A-4438 | Sequence 4438, Ap |
| 30 | 99.6 | 8.3 | 119 | 9 | US-09-864-761-19109 | Sequence 19109, A |
| 31 | 90 | 7.5 | 125 | 17 | US-10-242-535A-17878 | Sequence 17878, A |
| 32 | 90 | 7.5 | 125 | 18 | US-10-085-783A-17878 | Sequence 17878, A |
| 33 | 50 | 4.1 | 50 | 17 | US-10-131-827-2967 | Sequence 2967, Ap |
| 34 | 50 | 4.1 | 50 | 17 | US-10-131-827-3824 | Sequence 3824, Ap |
| 35 | 46.4 | 3.9 | 60 | 10 | US-09-908-375-7933 | Sequence 7933, Ap |
| 36 | 42.4 | 3.5 | 999 | 14 | US-10-184-644-434 | Sequence 434, App |
| 37 | 42.4 | 3.5 | 999 | 14 | US-10-175-169-88 | Sequence 88, Appl |
| 38 | 41.8 | 3.5 | 220895 | 19 | US-10-123-155-10 | Sequence 10, Appl |
| 39 | 41.6 | 3.5 | 594 | 14 | US-10-146-731-10 | Sequence 10, Appl |
| 40 | 41.6 | 3.5 | 594 | 15 | US-10-140-472-10 | Sequence 10, Appl |
| 41 | 41.6 | 3.5 | 594 | 15 | US-10-141-761-10 | Sequence 10, Appl |
| 42 | 41.6 | 3.5 | 594 | 16 | US-10-142-885-10 | Sequence 10, Appl |
| 43 | 41.6 | 3.5 | 594 | 16 | US-10-158-790-10 | Sequence 10, Appl |
| 44 | 41.6 | 3.5 | 594 | 16 | US-10-137-871-10 | Sequence 10, Appl |
| 45 | 41.6 | 3.5 | 594 | 17 | US-10-137-871-10 | Sequence 10, Appl |

ALIGNMENTS

RESULT 1
US-08-731-499-7
; Sequence 7, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 1..1205
; LOCATION: 1..1205
; OTHER INFORMATION: /note= "cDNA clone 1b4 for a serine
; OTHER INFORMATION: threonine kinase"
US-08-731-499-7

Query Match 99.7%; Score 1201; DB 8; Length 1205;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGCGGTGAGTCGCGCCCGCCAGTCAGTCAGCGCTGACTCGGGGGCTTCTCCACTATC 60
Db 1 GCGGCGGTGAGTCGCGCCCGCCAGTCAGTCAGCGCTGACTCGGGGGCTTCTCCACTATC 60

Qy 61 GCTTACCTACCTCCCTCTGCGAGAACCGCGGATATGCTCGCGTGGCCCGCGCGC 120
Db 61 GCTTACCTACCTCCCTCTGCGAGAACCGCGGATATGCTCGCGTGGCCCGCGCGC 120

Qy 121 ATTTCCTCTCCCGGTGCTTCCCTCTCTCGGGCTTCTGCTCTCTCGCTCCGATGGC 180
Db 121 ATTTCCTCTCCCGGTGCTTCCCTCTCTCGGGCTTCTGCTCTCTCGCTCCGATGGC 180

Qy 181 GGCAGCGCGTGCACACCAAGGCGCCCTTCCCTGGATACGGTGACTTCTACAGGTCA 240
Db 181 GGCAGCGCGTGCACACCAAGGCGCCCTTCCCTGGATACGGTGACTTCTACAGGTCA 240

Qy 241 TTCCCAAGCAAGTTCGCTGCTGAGTTCGACACCCAGTACCCCTACGGTGAGAACA 300
Db 241 TTCCCAAGCAAGTTCGCTGCTGAGTTCGACACCCAGTACCCCTACGGTGAGAACA 300

Qy 301 GGATGAGTTCAAGCGCTTCTGAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAG 360
Db 301 GGATGAGTTCAAGCGCTTCTGAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAG 360

Qy 361 GTGGGATCTCAGATTATGTGACAAAGCTGAAATGACAGCTGAGTGAGAAATACAGCTGG 420
Db 361 GTGGGATCTCAGATTATGTGACAAAGCTGAAATGACAGCTGAGTGAGAAATACAGCTGG 420

Qy 421 ACAAGAGAGTACCCATCTTCTACTCTCTCGGGATGGGACTTTGAGAACCCAGTCCC 480
Db 421 ACAAGAGAGTACCCATCTTCTACTCTCTCGGGATGGGACTTTGAGAACCCAGTCCC 480

Qy 481 ATACACTGGGCGAGTTAGGTTGGACCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTAC 540
Db 481 ATACACTGGGCGAGTTAGGTTGGACCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTAC 540

Qy 541 CTAGGTATGCTGCTGCTGCTGTATACGACGCCCTGGCGGGGAGTTCTATCAGGGCT 600
Db 541 CTAGGTATGCTGCTGCTGCTGTATACGACGCCCTGGCGGGGAGTTCTATCAGGGCT 600

Qy 601 CTGGTGTGAGGCGCCGAGGCCCTCTTGNAGCAGGGCAAGATACCTCTCAAGTGTCAA 660
Db 601 CTGGTGTGAGGCGCCGAGGCCCTCTTGNAGCAGGGCAAGATACCTCTCAAGTGTCAA 660

Qy 661 GGAGACTCAGAAGAGTGGGCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCA 720
```

```

Db 661 GGAGACTCAGAAGAGTGGGCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCA 720
Qy 721 GGGAGCACATTCAGAGCATCAGAGATGACACGGATCGCCAGGCTGATTTGAGAAACAGA 780
Db 721 GGGAGCACATTCAGAGCATCAGAGATGACACGGATCGCCAGGCTGATTTGAGAAACAGA 780
Qy 781 TGAGTGACGCGCAGAAAGAGGAGCTCCAGAAAGAGCTTAAACATCTCTGACTGCTTCCAGAA 840
Db 781 TGAGTGACGCGCAGAAAGAGGAGCTCCAGAAAGAGCTTAAACATCTCTGACTGCTTCCAGAA 840
Qy 841 GAAGGGGCGCGAGAAAGAGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTGGTG 900
Db 841 GAAGGGGCGCGAGAAAGAGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTGGTG 900
Qy 901 GGGGTAGGAGGGGANNAGTTAACTGCTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGY 960
Db 901 GGGGTAGGAGGGGANNAGTTAACTGCTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGY 960
Qy 961 MSKGGGAAAGWGTACTTAACCCACGATTTCTGAGCCCTGAGTATGCTTGGACATTTGATGC 1020
Db 961 MSKGGGAAAGWGTACTTAACCCACGATTTCTGAGCCCTGAGTATGCTTGGACATTTGATGC 1020
Qy 1021 TAACATGACCATGCTTGGGATGCTCTAGCTGCTGGGATAGCTGGAGCACATTACTCA 1080
Db 1021 TAACATGACCATGCTTGGGATGCTCTAGCTGCTGGGATAGCTGGAGCACATTACTCA 1080
Qy 1081 GGTGGCTGGTGAATGACACTCAGAGGAATGAGTCTATAGAGAGGAGAGAGGTGT 1140
Db 1081 GGTGGCTGGTGAATGACACTCAGAGGAATGAGTCTATAGAGAGGAGAGAGGTGT 1140
Qy 1141 ACTGCCAGGCTTTTGACAGATGTAATCTCAATTCAATTTAAAGTTTCAGTGTTTGGTTA 1200
Db 1141 ACTGCCAGGCTTTTGACAGATGTAATCTCAATTCAATTTAAAGTTTCAGTGTTTGGTTA 1200
Qy 1201 AGTGG 1205
Db 1201 AGTGG 1205

RESULT 2
US-09-925-301-566
; Sequence 566, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 566
; LENGTH: 1334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1253)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1307)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1309)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1312)
; OTHER INFORMATION: n equals a,t,g, or c
```

US-09-925-301-566

Query Match 88.2%; Score 1062.8; DB 9; Length 1334;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1193; Conservative 5; Mismatches 5; Indels 12; Gaps 12;

```
QY 1 GCGCGCTGAGTCGCGCCCCCCCCAGTCAGTGACCGCTGACTCGGGCGCTTCTCCACTATC 60
DB 40 GCGCGGTGAGTCGCGCCCCCCCCAGTCAGTGACCGCTGACTCGGGCGCTTCTCCACTATC 99
QY 61 GCTTACCTACCTCCCTCTGTCAGGAACCGGGCGATATGGCTGCGCTGTCGCCCGCGCGCG 120
DB 100 GCTTACCTACCTCCCTCTGTCAGGAACCGGGCGATATGGCTGCGCTGTCGCCCGCGCGCG 159
QY 121 ATTCTCTCCCGCTGCTT-CCCTCTCTCTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 179
DB 160 ATTCTCTCCCGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 219
QY 180 CGCGAGCGCTGCACACCAA-GGCGCCCTTCCCTGGATACGGTCACCTTCTACAAGT 238
DB 220 CGCGAGCGCTGCACACCAAAGGGCGCTTCCCTGGATACGGTCACCTTCTACAAGT 279
QY 239 CATTCCTCCAAAGCAAGTTGCTG- TGCTGAAGTTTCGACACCCAGTACCCCTACGGTGAGAA 297
DB 280 CATTCCTCCAAAGCAAGTTGCTG- TGCTGAAGTTTCGACACCCAGTACCCCTACGGTGAGAA 339
QY 298 GAGGATGAGTTCAAGCGTCTT-CTGAAAACCTCGGCTTCCAGCGATGATCTCTGGTGGC 356
DB 340 GAGGATGAGTTCAAGCGTCTTCTGTAACCTCGGCTTCCAGCGATGATCTCTGGTGGC 399
QY 357 AGAGTGGGGATCTCAGATTAT-GTGACAAGCTGAACATGGAGCTGAGTGAGAAATCAA 415
DB 400 AGAGTGGGGATCTCAGATTATGGTGACAAGCTGAACATGGAGCTGAGTGAGAAATCAA 459
QY 416 GCTGACAAAGAGACTACCCA-TCTTCTACCTCTTCCGGATGGGACTTTGAGAACCC 474
DB 460 GCTGACAAAGAGACTACCCAAGTCTTCTACCTCTTCCGGATGGGACTTTGAGAACCC 519
QY 475 AGTCCCATACACTGGGGAGTT-AGGTTGGAGCCATCCAGCGCTGGCTGGAAGGGCAAGG 533
DB 520 AGTCCCATACACTGGGGAGTTAAGTTGGAGCCATCCAGCGCTGGCTGGAAGGGCAAGG 579
QY 534 GGTCTACCTAGTATGCTGG-TGCTGCTGTATACGACGCTTGGCGGGGAGTTCTAT 592
DB 580 GGTCTACCTAGTATGCTGGTGGCTGCTGTATACGACGCTTGGCGGGGAGTTCTAT 639
QY 593 CAGGCGCTCTGTGTGGAGG-CGCGCAGGCTCTTTGAAGCAGGGGCAAGATAAAGCTCTC 651
DB 640 CAGGCGCTCTGTGTGGAGGCGCGCAGGCTCTTTGAAGCAGGGGCAAGATAAAGCTCTC 699
QY 652 AAGTGTGAAGGAGACTCAGAA-AGTGGGCGAGCAATACCTGAAGATCATGGGGAAGAT 710
DB 700 AAGTGTGAAGGAGACTCAGAAAGTGGGCGGAGCAATACCTGAAGATCATGGGGAAGAT 759
QY 711 CTTAGACCAAGGGAGCACTT-CCAGCATCAGAGATGACCGATCGCCAGGCTGATTGA 769
DB 760 CTTAGACCAAGGGAGCACTTCCAGCATCAGAGATGACCGATCGCCAGGCTGATTGA 819
QY 770 GAAGAACCAAGATGAGTGAC-GGCAGAGAGGAGCTCCAGAGAGCTTAAACATCTGCAC 828
DB 820 GAAGAACCAAGATGAGTGACCGGAAGAGAGGAGCTCCAGAGAGCTTAAACATCTGCAC 879
QY 829 TGCCTTCCAGAAAGAGGGGGCCGAGAAAGAGGAGCTGTAAAGGCTGTGTGTGATTTTC 888
DB 880 TGCCTTCCAGAAAGAGGGGGCCGAGAAAGAGGAGCTGTAAAGGCTGTGTGTGATTTTC 939
QY 889 CAGGGTTTGGTGGGGTAGGGAGGGGAGNAGTTAACTGCTGTGTGTGANTCCCTTGTGA 948
DB 940 CAGGGTTTGGTGGGGTAGGGAGGGGAGAGTTAACTGCTGTGTGTGANTCCCTTGTGA 999
QY 949 ATATAAGGGGGYMSKGGGAAGAGWGTACTTAACCCAGATTTCTGAGCCCTGAGTATGCT 1008
DB 1000 ATATAAGGGGGTAGTGGGGAAGAGTGTACTTAACCCAGATTTCTGAGCCCTGAGTATGCT 1059
```

RESULT 3

```
US-10-205-823-43
; Sequence 43, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-43
```

```
Query Match 66.7%; Score 803.8; DB 15; Length 963;
Best Local Similarity 97.8%; Pred. No. 2,6e-253;
Matches 934; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

QY 87 CCGCGATATGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 145
DB 4 CCGCGATATGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 63
QY 146 TCCTGGGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
DB 64 TCCTGGGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 123
QY 205 CCGCTTCCCTCGATACGGTCACTTCTTACAAGGTCACTTCCCAAGAGCAAGTTTCGTC-TGG 263
DB 124 CCGTTCCTCGATACGGTCACTTCTTACAAGGTCACTTCCCAAGAGCAAGTTTCGTC-TGG 183
```


Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 273
; LENGTH: 4529
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_053961
US-10-191-803-273

Query Match 42.9%; Score 517.2; DB 17; Length 4529;
Best Local Similarity 78.5%; Pred. No. 1.2e-158;
Matches 832; Conservative 1; Mismatches 201; Indels 26; Gaps 17;

QY 1 GCGCGCGTGAAGTCCGCCCCCAGTCACTGACCGCTGACTCGGGCGTCTTCCACTATC 60
Db 3277 GCGCGCGGAGAGCGCCCTCCAGTGAAGTGAACCGCCACTCGGAGCGTC-CTGCTGA 3335
QY 61 GCTTACCTACCTCCTCTGAGGAACCGGCGATATGGTGGTGGTGGTGGTGGTGGTGG 120
Db 3336 GCTTGGCCATCTCCCGCTCCCAACATCGGCGGTGATGGTGGTGGTGGTGGTGGTGG 3395
QY 121 ATTTCTCTCCCGCTGCT-TCCTTCTCTCTGGGCTTCTGCTCTCTCGCTCCGCTGG 179
Db 3396 TTGCTCTCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3455
QY 180 CGGACGCGCTGCACACAA-GGCGCCCTTCCCTGGATACGTCACCTTCTACAAGGT 238
Db 3456 CGGACGCGCTGCACACGAGGCGCCCTTCCCTGGACACAGTCACCTTCTACAAGGT 3515
QY 239 CATTCCTCCAAAGCAAGTTCGTC-TGGTGAAGTTCGACACCCAGTACCCCTACGGTGAA 297
Db 3516 CATTCCTCCAAAGCAAGTTCGTC-TGGTGAAGTTCGACACCCAGTACCCCTATGAGAGAA 3575
QY 298 CGAGATGAGTTCAGCGCTCT-TCGAAACTCGGCTTCGACGCGATGATCTTGGTGGC 356
Db 3576 CGAAGTGAAGTTCAGCGCTCTGCTGAGAACTCAGCGCTTCGACGCGATGATCTTGGTGGC 3635
QY 357 AGAGTGGGATCTCAGATTAT-GTGACAACTGAACTGAGCTGAGTGAGAAATACAA 415
Db 3636 AGAGTGGGATCTCAGACTATGTGACAACTGAACTGAGCTGAGTGAGAGTACAA 3695
QY 416 GCTGACAAAGAGACTACCCA-TCTTCTACCTTCCGGGATGGGACTTTGAGAACCC 474
Db 3696 GCTGACAAAGAGACTACCCAGTCTTCTACCTTCCGGGATGGGACTTTGAGAACCC 3755
QY 475 AGTCCCATACCTGGGCGAGTT-AGTTGGAGCCATCCAGCGCTGGCTGAAGGGCAGG 533
Db 3756 TGTCCTACACAGCGGGGAGTTAAAGTTGGAGCCATCCAGCGCTGGCTCAAAGGGCAGG 3815
QY 534 GGTCTACCTAGTATGCTGG-TGCCTGCTCTATACGACGCGCTGGCGGGGAGTTTCA 592
Db 3816 AGTCTATCTGGGATGCTTGGATGCTGCTGCTGATGATGCTGCTGGCGGCGGCTTCA 3875

QY 593 CAGGCGCTCTGTTGGAGG-CCGCCAGGCGCTCTTGAAGCAGGCGGCAAGATAACTCTC 651
Db 3876 CGAGCGCTCCAGCAGAGAGAGCGCCCGCCAGGCCATCTCTGAACAGGCGGCAAGATGGGCTCTC 3935
QY 652 AAGTGTGAAGGAGACTCAGAAG-AGTGGCGCGAGCAATACCTGAAGATCATGGGGAAGAT 710
Db 3936 AGTGTGAAGGAGACACAGAGAGTGGCCAGTCAGTCACTCAAGATCATGGGGAAGAT 3995
QY 711 CTTAGACCAAGGGGAGCACTT-CCAGCATCAGAGATGACAGGATCGCGAGGCTGATGA 769
Db 3996 CTTGACCAAGGTGAAGACTTCCCGGCTCCGAGCTGGCCCGGATCAGTAAGCTCATTTGA 4055
QY 770 GAAGAACAGATGAGTGACGCGCAGAGAGAGGCTCCAGAGAGCTTAAACATCTGACT 829
Db 4056 GAACAAGATGAGTGAGG--GTAAGAGGAAGAGCTGCAGAGAGCCTCAACATCTCACC 4113
QY 830 GCCTTCCAGAAAGAGGGGCGCCGAGAAAGAGAGCTGTAAAGAGGCTGTGTGATTTTCC 889
Db 4114 GCCTTCCGCAAGAAAGGCGCCGAGAGAGAGAGCTGT-----GAGGCGACCGAGCTGT 4167
QY 890 AGGCTTTGCTGGGTAGGGAGGGGAGGAGTTAACTGCTGCTGTGANTCCCTTGTGAA 949
Db 4168 AGGCTTTGCTGAGGCGGGGAGGG--AGAGCGGCTGCTGACTGTGGGCGCTGTGG-- 4223
QY 950 TATAAGGGGYSKGGGNAAGWGTACTAACCCAGGATCTGAGCCCTGAGTATGCCCTG 1009
Db 4224 TGAAGGGGCGAGTGGAGCAAGAGCTCTGAGCCAGAGACCTGTGCCCCCGAG--TGCCTG 4281
QY 1010 GACATTGATGCTAAACATGACCATGCTTGGGATGTCTCTAG 1049
Db 4282 GACACTGAGGCTGTGAGACCATACCTTGGACATCTTGG 4321

RESULT 6

US-10-152-319A-2052
; Sequence 2052, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2052

```
; LENGTH: 4529
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_053961
US-10-152-319A-2052

Query Match      42.9%; Score 517.2; DB 18; Length 4529;
Best Local Similarity 78.5%; Pred. No. 1.2e-158;
Matches 832; Conservative 1; Mismatches 201; Indels 26; Gaps 17;

QY 1 GCGCCGCTGAGTCCCGCCCCCAGTCACGTGACCCGCTGACTCGGGCGTTCTTCCATATC 60
Db 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3277 GCGCCGCGGGAGCGGCCCTCCAGTGAGCTGACCCGCCACTCGAGCGCTC-CTGCTGTA 3335
Db 3277 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GCTTACCTACCTCCTCTCGAGAACCCGGCGATATGCTGCGCTGTCGCCCGCGCGC 120
Db 61 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3336 GCTTCCCATCTTCCCGCTCCCAATCCGCGGTATGCTGCGCGCTTCTTGGGCGCGT 3395
Db 3336 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 ATTCTCTCCCGCTGCT-TCCCTTCTCTGCGCTTCTGCTCTCTCTCCGCTCCGCATGG 179
Db 121 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3396 TTCGCTCTCCCACTGCTGCTGCTTCTCTTGGCCCTGCTCTCTCTCTGCTCTCATGG 3455
Db 3396 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 180 CGGACAGCGGCTGACACAA-GGCGCCCTTCCCTCGATACGGTCACCTTCTTCAAGGT 238
Db 180 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3456 CGCCAGCGGCTGACACGAGAGGCGCCCTTCCCTTGGACACAGTCACCTTCTTCAAGGT 3515
Db 3456 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 239 CATTCCTCCAAAGCAAGTTGCTC-TGCTGAAGTTGACACCCAGTACCCCTACGGTGAGAA 297
Db 239 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3516 CATTCCTCCAAAGCAAGTTGCTC-TGCTGAAGTTGACACCCAGTACCCCTATGAGAGAA 3575
Db 3516 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 298 CGAGATGAGTTCAAGCGTCT-TCTGAAACTCGGCTTCCAGCGATGATCTCTTGGTGGC 356
Db 298 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3576 GCAAGATGAGTTAAGCGTCTGGCTGAGACTCAGCTCCAGCGATGATCTCTTGGTGGC 3635
Db 3576 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 357 AGAGTGGGATCTCAGATTAT-GTGACAGCTGAACATGAGCTGAGTGAGAAATACAA 415
Db 357 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3636 AGAGTGGGATCTCAGACTATGCTGACAACTGAACATGAGCTGAGTGAGAAATACAA 3695
Db 3636 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 416 GCTGACAAAGAGACTACCCA-TCTTCTACCTCTTCCGGGATGGGACTTTGACAAACC 474
Db 416 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3696 GCTGACAAAGAGACTACCCAGTCTTCTACCTCTTCCGGGATGGGACTTTGAGAAATCC 3755
Db 3696 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 475 AGTCCCATACATCGGGGAGTT-AGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGG 533
Db 475 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3756 TGTCCTATACAGGGGCGAGTTAAAGTTGAGCCATCCAGCGCTGGCTCAAGGGGCGAGG 3815
Db 3756 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 534 GGTCTACCTAGTATGCTCTGG-TGCTGCTGTATACGACGCTTGGCGGGGAGTTTCAAT 592
Db 534 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3816 AGTCTATCTGGCATGCTTGGATGTCTGCTCGTACGATGCTTGGCGGGGCGAGTTTCAAT 3875
Db 3816 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 593 CAGGCGCTCTGTGTGGAGG-CGCGCAGGCCCTTGAAGCAGGGGCAAGATAACCTCTC 651
Db 593 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3876 CAGGCGCTCCAGCAGAGAGGCCCGCAGGCCATCTGAAACAGGGGCGAGGATGGCCTCTC 3935
Db 3876 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 652 AAGTGTGAAGAGACTCAGAAG-AGTGGGCCGAGCAATACCTGAAGATCATGGGGAAGAT 710
Db 652 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3936 AGGTGTGAAGAGACAGACAGAAGTGGCCAGTCAGTACCTGAAGATCATGGGGAAGAT 3995
Db 3936 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 711 CTTAGACCAAGGGGAGCACTT-CCAGCATCAGAGATGACAGGATCGCCAGGCTGATTGA 769
Db 711 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3996 CTTGACCAAGGTGAAGACTTCCCGGCTCCGAGCTGGCCCGGATCAGTAAGCTCATTTGA 4055
Db 3996 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 770 GAAGACAGATGAGTGAAGCGGACAGAGGAGGAGCTCCAGAGAGCTTAAACATCTCTGACT 829
Db 770 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4056 GAACAGATGAGTGAAG--GTAAAGAGAGAGCTGACGAGAGGCTCAACATCTCTACC 4113
Db 4056 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 830 GCCTTCCAGAAAGGGGCCGAGAAAGAGAGCTGTAAAAAGGCTGTCTGTGATTTTCC 889
Db 830 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4114 GCCTTCCGCAAGNAGGCCCGAGAGAGGAGCTGT-----GAGGGCACCAGGCTGT 4167
Db 4114 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 890 AGGTTTGTGGGGTAGGGAGGGGAGNAGTTAACTGCTGGCTGTGANTCCCTTGTGGAA 949
Db 890 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 4168 AGGTTTGTGAGGGCGGGGAGGG--AGAGCCGCTGCTGACTGTGGGGCCCTGTGG-- 4223
QY 950 TATAAGGGGGYMSKGGGAAAGWGTACTATAACCCACGATTTCTGAGCCCTGAGTATGCTG 1009
Db 4224 TGGAGGGGGCAGTGGAGCAAGACAGTCTCTGAGCCAGAGACCTGTGCCCCGAG--TGCTG 4281
QY 1010 GACATTGATGCTAATACATGACCATGCTTTGGGATGCTCTAG 1049
Db 4282 GACACTGAGGCTGTGAGACCATTACCTCTGGACATCTCTGG 4321

RESULT 7
US-10-956-157-6905
; Sequence 6905, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6905
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-6905

Query Match      42.8%; Score 515.8; DB 21; Length 600;
Best Local Similarity 97.5%; Pred. No. 1.3e-158;
Matches 580; Conservative 5; Mismatches 4; Indels 6; Gaps 6;

QY 441 TCTACCTCTTCCGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGCGAGTT-AGG 499
Db 1 TCTACCTCTTCCGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGCGAGTTAAGG 60
QY 500 TTGAGGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACTAGGTATGCTTG-TCGC 558
Db 61 TTGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACTAGGTATGCTTGTTGCC 120
QY 559 TGCCTGTATACGACGCCCTGGCCGGGGAGTTTCATCAGGGCTCTGGTGTGAGG-CCGCC 617
Db 121 TGCCTGTATACGACGCCCTGGCCGGGGAGTTTCATCAGGGCTCTGGTGTGAGGCCGCC 180
QY 618 AGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAG-AGT 676
Db 181 AGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAGAAAGT 240
QY 677 GGGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAG 735
Db 241 GGGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGGACTTCCCAG 300
QY 736 CATCAGATGACACGGATCCCGAGTTCAGTGTGAGAGAAACAAGATGATGACGCG-AGA 794
Db 301 CATCAGATGACACGGATCCCGAGTTCAGTGTGAGAGAAACAAGATGATGACGCGCAAGA 360
QY 795 AGGAGGAGCTCCAGAGAGCTTAAACATCTCCCTGACTGCTCCAGAGAGGGGCGGAGA 854
Db 361 AGGAGGAGCTCCAGAGAGCTTAAACATCTCCCTGACTGCTCCAGAGAGGGGCGGAGA 420
QY 855 AAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAGGGG 914
Db 421 AAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAGGGG 480
QY 915 ANAGTTAACCTGCTGGCTGTGANTCCCTTGTGGAATATTAAGGGGYSKGGGAAAGGG 974
Db 481 AGAGTTAACCTGCTGGCTGTGAGTCCCTTGTGGAATATTAAGGGGCTAGTGGGAAACAGTGG 540
QY 975 TACTAACCCAGATTCAGCGCTGAGTATCCCTGGACATTCATGCTAACATGAC 1029
Db 975 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 541 TACTAACCCACGATCTGAGCCCTGAGTATGCTGACATTTGATGCTAACATGAC 595

RESULT 8
US-10-388-934-523
; Sequence 523, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 523
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)
US-10-388-934-523

Query Match 41.2%; Score 496.6; DB 17; Length 1139;
Best Local Similarity 78.8%; Pred. No. 3.8e-152;
Matches 792; Conservative 1; Mismatches 187; Indels 25; Gaps 16;

QY 56 CTATCGTTACTACTCTCCCTCTGCAGNACCGGGGATATGGCTGCGCTGTGCGCCGC 115
DB 5 CTGTAGTTCGCCCATCTCCCGTCCCAATCCGGCGTGANGGCTGCGCGCTTCTCTGC 64

QY 116 GCGCATTTCTCCCGCTGTCT-TCCCTTCTCTGGGCTTCTGTCTCTCTCCGCTCG 174
DB 65 GCGTTTCTCCCGCTGTCTGCGTCTCTCTCTGGCTGTCTCTCTCTCTCTCTCTCT 124

QY 175 CATGCGGCGAGCGGCTGCACACCAA-6GCGCCCTTCCCTCGATACCGTCACTTTCTAC 233
DB 125 CATGCGGCGAGCGGCTGCACACGAAGGCGCGCTTCCCTTGACACAGTCACTTTCTAC 184

QY 234 AGGTGATCTCCCAAGCAAGTTCGTC-TGGTGAAGTTCGACACCCAGTACCCCTACGGT 292
DB 185 AAGGTGATCTCCCAAGCAAGTTCGTC-TGGTGAAGTTCGACACCCAGTACCCCTATGGA 244

QY 293 GAGAGCAGGATGAGTTCAAGGCTCT-TCTGAAACTCGGCTTCCAGCGATGATCTCTTG 351
DB 245 GAGAGCAGGATGAGTTTAAAGCTCTGGCTGAGACTAGCTTCCAGCGATGATCTCTTG 304

QY 352 GTGGCAGAGGTGGGATCTCAGATTAT-GTGAACAAGCTGAAATGAGCTGAGTGAGAAA 410
DB 305 GTGGCAGAGGTGGGATCTCAGACTATGTGACACAGCTGAAATGAGCTGAGTGAGAG 364

QY 411 TACAGCTGGAACAAGAGAGTACCCA-TCTTCTACCTTCCGGGATGGGAGCTTTGAG 469
DB 365 TACAGCTGGAACAAGAGAGTACCCA-TCTTCTACCTTCCGGGATGGGAGCTTTGAG 424

QY 470 RACCCAGTCCCATACACTGGGCGAGTT-AGGTTGAGCCATCCAGGCTGGCTGAAGGG 528
DB 425 AATCTGTGCCCATACAGCGGGGAGTTAAAGTTGAGCCATCCAGCGCTGGCTCAAGGG 484

QY 529 CAAGGGGTCTACTAGGTATGCTGG-TGCCCTGCTGTATACGAGCCCTGCGCGGGAG 587
DB 485 CAGGGAGTCTATCTGGGATGCTGGATGTCTGCTTGGCTAGATGCTTGGCGGGCAG 544

QY 588 TTCATCAGGCTCTGTGTGTGAGG-CGCGCAGGCGCTTCTTGAAGCAGGGGCAAGATAAC 646
DB 545 TTCATCAGGCTCTCAGCAGAGAGGCGCGCCAGGCGCATCTTGAACAGGGGCGAGTGGC 604

QY 647 CTCTCAAGTGTGAAGGACTCAGAG-AGTGGGCGGAGCANTACTGAGATCATGGG 705

Db 605 CTCTCAGGTGTGAAGGAGACACAAGAGTGGCGCAGTCACTACTGAAGATCATGGGG 664

QY 706 AAGATCTTTAGACCAAGGGGAGCATT-CCAGCATCAGAGATGACACGGATCCGAGGCTG 764
DB 665 AAGATCTTTGGACCAAGGTGAAGACTTCCCGGCTCCGAGCTGGCCGGATCAGTAAGCTC 724

QY 765 ATTGAGAAGAACAGATGAGTGACCGCAGAGGAGGAGCTCCAGAGAGCTTTAAACATCC 824
DB 725 ATTGAGAAGAACAGATGAGTGAGG--GTAAGAAGAAAGAGCTGCAGAGGAGCCTCAACATCC 782

QY 825 TGACTGCTCTCCAGAAAGGGGCGGAGAAAGAGAGCTGTAAAGAGCTGTCTGTGAT 884
DB 783 TCACCGCTTCCGCAAGAAAGCGCGGAGAGGAGCTGT-----GAGGCGACACAG 836

QY 885 TTTCCAGGTTTGTGGGGTAGGAGGGGAGNAGTTAACCTGCTGGCTGTGTGANTCCCTTG 944
DB 837 CTGCTAGGTTTGTCTGAGGGGGGAGGG--AGAGCGGCTGTGACTGTGCGGCGCTGT 894

QY 945 TGGATATTAAGGGGYSKGGGAAAGGTTACTAACCCAGATTTCTGAGCCCTGAGTAT 1004
DB 895 GGG--TGGAGGGGGCAGTGGAGCAAGCAGTCTTGAGCCAGAGACTGTGCCCCGAG--T 950

QY 1005 GCCTGGACATTCATCTAACATGACCATGCTTGGGATGTCTCTAG 1049
DB 951 GCCTGGACATGAGGCTGCTGAGACCATAACCTTGACATCTCTGG 995

RESULT 9
US-10-242-535A-37352
; Sequence 37352, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37352
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (24)..(25)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-37352

Query Match 33.6%; Score 404.8; DB 17; Length 438;
Best Local Similarity 97.0%; Pred. No. 4e-122;
Matches 425; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

QY 691 CTGAAGATCATGGGAGAGATCTTAGACCAAGGGGAGCACTT-CCAGCATCAGAGATGACA 749
DB 1 CTGAAGATCATGGGAGAGATCTTNNACCAAGGGGAGGACTTCCCAGCATCAGAGATGACA 60

QY 750 CGATCGCCAGCGTGAATTGAGAAGAACAAAGATGATGAC-GGCAGAGAGGAGGCTCCAG 808
DB 61 CGATCGCCAGCGTGAATTGAGAAGAACAAAGATGATGACGCGGAGAGAGGAGCTCCAG 120

QY 809 AAGAGCTTAAACATCTGACTGCTTCCAGAGAGGGGGCGAGAAAGAGGAGCTGTAA 868
DB 121 AAGAGCTTAAACATCTGACTGCTTCCAGAGAGAGGGGGCGAGAAAGAGGAGCTGTAA 180

```
Qy 869 AAAGCTCTCTGATTTTCCAGGGTTGGTGGGGTAGGGAGGAGNAGTTAACCTGCT 928
Db 181 AAAGCTCTCTGATTTTCCAGGGTTGGTGGGGTAGGGAGGAGTTAACCTGCT 240
Qy 929 GGCTGTGANTCCCTTGTGGAATATAAGGGGYSKGGGAAAGWGGTACTAACCCACGAT 988
Db 241 GGCTGTGAGTCCCTTGTGGAATATAAGGGGAGTAGTGGGAAAGTGGTACTAACCCACGAT 300
Qy 989 TCTGAGCCCTCAGTATGCTCGACATTCATCTACATGACCATGCTTGGGATGCTCTTA 1048
Db 301 TCTGAGCCCTCAGTATGCTCGACATTCATCTACATGACCATGCTTGGGATGCTCTTA 360
Qy 1049 GCTGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGAATACACACCTCAGAAG 1108
Db 361 GCTGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGAATACACACCTCAGAAG 420
Qy 1109 GAATGAGTCTATAGAGA 1126
Db 421 GAATGAGTCTATAGAGA 438

RESULT 10
US-10-085-783A-37352
; Sequence 37352, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37352
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (24)..(25)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-37352

Query Match 33.6%; Score 404.8; DB 18; Length 438;
Best Local Similarity 97.0%; Pred. No. 4e-122;
Matches 425; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

Qy 691 CTGAAGATCATGGGAAGATCTTACACCAAGGGAGGACCTT-CCAGCATCAGAGATGACA 749
Db 1 CTGAAGATCATGGGAAGATCTTNNACCAAGGGAGGACTTCCAGCATCAGAGATGACA 60
Qy 750 CGGATCGCCAGCTGATTGAGAAGAACAAAGATGATGAC-GGCAGAGAGAGAGCTCCAG 808
Db 61 CGGATCGCCAGCTGATTGAGAAGAACAAAGATGATGACGCGGAAGAGAGAGCTCCAG 120
Qy 809 AAGAGCTTAAACATCTGACTCCCTTCCAGAGAGAGGGCCGAGAAAGAGAGCTGTAA 868
Db 121 AAGAGCTTAAACATCTGACTCCCTTCCAGAGAGAGGGCCGAGAAAGAGAGCTGTAA 180
Qy 869 AAAGCTCTCTGATTTTCCAGGGTTGGTGGGGTAGGGAGGAGNAGTTAACCTGCT 928
Db 181 AAAGCTCTCTGATTTTCCAGGGTTGGTGGGGTAGGGAGGAGTTAACCTGCT 240
Qy 929 GGCTGTGANTCCCTTGTGGAATATAAGGGGYSKGGGAAAGWGGTACTAACCCACGAT 988
Db 241 GGCTGTGAGTCCCTTGTGGAATATAAGGGGAGTAGTGGGAAAGTGGTACTAACCCACGAT 988
```

```
Db 241 GGCTGTGAGTCCCTTGTGGAATATAAGGGGAGTAGTGGAAAGAGTGTACTAACCCACGAT 300
Qy 989 TCTGAGCCCTCAGTATGCTCGACATTCATCTACATGACCATGCTTGGGATGCTCTTA 1048
Db 301 TCTGAGCCCTCAGTATGCTCGGACATTCATCTACATGACCATGCTTGGGATGCTCTTA 360
Qy 1049 GCTGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGAATACACACCTCAGAAG 1108
Db 361 GCTGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGAATACACACCTCAGAAG 420
Qy 1109 GAATGAGTCTATAGAGA 1126
Db 421 GAATGAGTCTATAGAGA 438

RESULT 11
US-10-066-543-1394
; Sequence 1394, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121,563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1394
; LENGTH: 395
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 167, 308, 353
; OTHER INFORMATION: n = A,T,C or G
US-10-066-543-1394

Query Match 27.3%; Score 328.8; DB 14; Length 395;
Best Local Similarity 97.5%; Pred. No. 4e-97;
Matches 385; Conservative 0; Mismatches 5; Indels 5; Gaps 5;

Qy 506 CCATCCAGCGCTGGCTGAAGGGCAAGGGTCTACCTAGGTATGCCTGG-TGCCTGCCTG 564
Db 1 CCATCCAGCGCTGGCTGAAGGGCAAGGGTCTACCTAGGTATGCCTGGTGCCTGCCTG 60
Qy 565 TATACGACGCCCTGGCCGGGAGTTTCATCAGGGCTCTGGTGTGGAGG-CGCCCCAGGCC 623
Db 61 TATACGACGCCCTGGCCGGGAGTTTCATCAGGGCTCTGGTGTGGAGGCCCGCCAGGCC 120
Qy 624 TCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAG-AGTGGGCGG 682
Db 121 TCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCANAAGAAGTGGGCGG 180
Qy 683 AGCAATACCTCAAGTATCATGGGAGATCTTAGACCAGGGGAGCATT-CCAGCATCAG 741
Db 181 AGCAATACCTCAAGTATCATGGGAGATCTTAGACCAGGGGAGACTTCCAGCATCAG 240
Qy 742 AGATGACACGGATCGCCAGGCTGATTGAGAAGAACAAAGATGATGAC-GGCAGAGAGGAG 800
Db 241 AGATGACACGGATCGCCAGGCTGATTGAGAAGAACAAAGATGATGACCGGGAAGAGGAG 300
Qy 801 AGCTCCAGAGAGCTTAAACATTCCTGACTGCTCCAGAGAGAGGGGCCGAGAAAGAGG 860
```


Db 301 AGCTCCANAGAGCTTAACATCTGAGTCCCTTCAGAGAGAGGGGCCGNAAGAGG 360
QY 861 AGCTGTAAAGAGCTGTCTGATTTTCCAGGTT 895
Db 361 AGCTGTAAAGAGCTGTCTGATTTTCCAGGTT 395

RESULT 12
US-09-918-995-5850
; Sequence 5850, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5850
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-5850

Query Match 26.6%; Score 321; DB 10; Length 405;
Best Local Similarity 98.5%; Pred. No. 1.5e-94;
Matches 387; Conservative 0; Mismatches 0; Indels 6; Gaps 6;
QY 88 CGGCGATATGGTCCGCTGTGTCGCCCGCGCATTTCTCCCCGCTGCTT-CCTTCT 146
Db 13 CGGCGATATGGTCCGCTGTGTCGCCCGCGCATTTCTCCCCGCTGCTTCCCTTCT 72
QY 147 CTTGGCTTCTGCTCTCTCCGCTCCGATGCGCGGCGCGCCCTGCACCAA-GGCGC 205
Db 73 CTTGGCTTCTGCTCTCTCCGCTCCGATGCGCGGCGCGCCCTGCACCAAAGGGGC 132
QY 206 CTTCCCTTGGATAGCTGCTACCTTCTACAGGTCTATCCCAAAGCAAGTTGTC-TGGT 264
Db 133 CTTCCCTTGGATAGCTGCTACCTTCTACAGGTCTATCCCAAAGCAAGTTGCTTGGT 192
QY 265 GAAGTTGCACACCCAGTACCTTCTACGCTGAGAGCAGGATGAGTTCAAGCGTCTT-CTGA 323
Db 193 GAAGTTGCACACCCAGTACCTTCTACGCTGAGAGCAGGATGAGTTCAAGCGTCTTGTGA 252
QY 324 AAATCTGGCTTCCAGGATGATCTCTTGGTGCAGAGTGGGATCTCAGATTAT-GTGA 382
Db 253 AAATCTGGCTTCCAGGATGATCTCTTGGTGCAGAGTGGGATCTCAGATTATGTGA 312
QY 383 CAAGCTGAACATGAGCTGAGTGAGAAATACAACTGGCAAGAGAGCTACCCA-TCCT 441
Db 313 CAAGCTGAACATGAGCTGAGTGAGAAATACAACTGGCAAGAGAGCTACCCAAGTCTT 372
QY 442 CTACCTTTCGGGATGGGATCTTTCAGAACCC 474
Db 373 CTACCTTTCGGGATGGGATCTTTCAGAACCC 405

RESULT 13
US-09-864-761-2375
; Sequence 2375, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2375
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109984.5
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 24
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
US-09-864-761-2375

Query Match 23.2%; Score 279.2; DB 9; Length 450;
Best Local Similarity 84.0%; Pred. No. 9.1e-81;
Matches 377; Conservative 0; Mismatches 59; Indels 13; Gaps 5;
QY 488 GGGGCGAGTTAGGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGCTTACCTTAGGTA 547
Db 5 GGGGCGAGTTAGGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGCTTACCTTAGGTA 64
QY 548 T-----GCCTGGTGCCTGCTGTATACAGCGCCCTGGCGGGAGTTCATCAGGGCC 599
Db 65 TGCCAGGTGGGCTGGTGGCCAGCTGCATAGGACACCTGGCC-AGGAGTTTCATCAGGGCC 123
QY 600 TCTGGTGGAGG-CCGCGAGCGCCCTTGTGAAGCGGGGCAAGATAACCTCTCAAGTGTG 658

Db 124 TCTGCGTGGAGCCCGCTAGGCCCTCTGGAAGCAGCGCAGGACCACTCTCAAGTGTG 183
Qy 659 AAGGAGACTCAGAAG-AGTGGCCGAGCAATACCTGGAAGATCATGGGGAAGATCTTAGAC 717
Db 184 AAGGAGACTGAGAAGACTGGCCGAGCAATACCTGGAAGATCACAGGGAAGATCTTAGAC 243
Qy 718 CAAGGGAGCACTTCCAGCATCAGAGATGACACCGATCCGAGGCTGATTTGAGAAGACA 777
Db 244 TAGGCAGAGGACTTCCAGCATAGATGACATGGATCTCCAGGCTGATTGAGAACAAGT 303
Qy 778 AGATGAGTGCAGCAGAGAGAGAGCTCCAGAGAGCTTAAACATCTCGATGCTGCTTCCA 837
Db 304 TCAGTGACG--GGAAGAAGGAGGAGCTCCAGAGAGCTTAAACATCTCGATGCTTCCA 361
Qy 838 GAAGAAGGGGCCGAGAGAGGAGCTCTFAAAGGCTGTCTGTGATTTTCCAGGGTTG 897
Db 362 GAAGAAGGGGCTGAGAAGCAGGAGTGTFAAAGGCTCTCTATGGTTTTCCAGGGTTG 421
Qy 898 GTGGGGGTAGGAGGGGAGTAACTTG 926
Db 422 GTAGGGGTGGGAGGGAGAGTTAACTTG 450

RESULT 14

US-09-918-18955
; Sequence 18955, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18955
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-18955

Query Match 22.6%; Score 271.8; DB 10; Length 345;
Best Local Similarity 97.9%; Pred. No. 2.2e-78;
Matches 328; Conservative 0; Mismatches 2; Indels 5; Gaps 5;
Qy 318 TTCTGAAAACCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGGATCTCAGATT 377
Db 11 TGCTGAAAACCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGGATCTCAAATTA 70
Qy 378 T-GTGAAGCTGAACATGAGCTGAGTGAGAAATACAACTGGACAAAGAGAGCTACCC 436
Db 71 TGGTGACAAGCTGAACATGAGCTGAGTGAGAAATACAACTGGACAAAGAGAGCTACCC 130
Qy 437 A-TCTTCTTACCTTCCGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGT 495
Db 131 AGTCTTCTTACCTTCCGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGT 190
Qy 496 T-AGGTTGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACCTAGGTATGCTGG 554
Db 191 TAAGGTTGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACCTAGGTATGCTGG 250
Qy 555 -TGCTGCTGTATACGAGCCCTGGCCGGGAGTTATCAGGGGCTCTGTGTGGAGG- 612
Db 251 TTGCTGCTGTATACGAGCCCTGGCCGGGAGTTATCAGGGGCTCTGTGTGGAGG- 310
Qy 613 CGCCAGGCGCTCTTGAAGCAGGGGCAAGATAACC 647
Db 311 CGCCAGGCGCTCTTGAAGCAGGGGCAAGATAACC 345

RESULT 15

US-10-242-535A-31517
; Sequence 31517, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31517
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (22)..(22)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-31517

Query Match 20.9%; Score 252.4; DB 17; Length 255;
Best Local Similarity 98.0%; Pred. No. 4.6e-72;
Matches 250; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Qy 941 CTTGTGGAATATAAGGGGYNKGGGAAAGWGTACTTAACCCAGATTCTGAGCCCTGA 1000
Db 1 CTTGTGGAATATAAGGGGNTANTGGAAAGTGGTACTTAACCCAGATTCTGAGCCCTGA 60
Qy 1001 GTATGCTGGACATGATGATTAACATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGG 1060
Db 61 GTATGCTGGACATGATGATTAACATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGG 120
Qy 1061 ATAGCTGGAGCACTTACTCTCAGGTGGCTGGTGAATGACACCTCAGAAGAAATGAGTGCTA 1120
Db 121 ATAGCTGGAGCACTTACTCTCAGGTGGCTGGTGAATGACACCTCAGAAGAAATGAGTGCTA 180
Qy 1121 TAGAGAGAGAGAGAGTGTACTGCCAGGTCTTTGACAGATGTAATTTCTCATTTCAATTA 1180
Db 181 TAGAGAGAGAGAGAGTGTACTGCCAGGTCTTTGACAGATGTAATTTCTCATTTCAATTA 240
Qy 1181 AAGTTTCAGTGTGTTT 1195
Db 241 AAGTTTCAGTGTGTTT 255

Search completed: July 4, 2005, 11:51:00
Job time : 705.092 secs

| Result No. | Score | Query | | | DB | ID | Description |
|------------|--------|-------|--------|---|----------|-------------|-------------|
| | | Match | Length | % | | | |
| 1 | 1034.8 | 85.9 | 1203 | 3 | CR625433 | full-length | |
| 2 | 1000.8 | 83.1 | 1152 | 3 | CR600542 | full-length | |
| 3 | 996.8 | 82.7 | 1375 | 3 | CR623789 | full-length | |
| 4 | 987.8 | 82.0 | 1139 | 3 | CR617833 | full-length | |
| 5 | 967.8 | 80.3 | 1119 | 3 | CR602804 | full-length | |
| 6 | 967.8 | 80.3 | 1119 | 3 | CR611287 | full-length | |
| 7 | 959.8 | 79.7 | 1111 | 3 | CR604444 | full-length | |
| 8 | 959.8 | 79.7 | 1111 | 3 | CR617687 | full-length | |
| 9 | 958.8 | 79.6 | 1110 | 3 | CR603906 | full-length | |
| 10 | 928.4 | 77.0 | 1137 | 5 | CR603042 | full-length | |
| 11 | 910.8 | 75.6 | 1058 | 3 | CR602432 | full-length | |
| 12 | 906.8 | 75.3 | 1062 | 3 | CR602432 | full-length | |
| 13 | 904.4 | 75.1 | 1098 | 3 | CR594614 | full-length | |
| 14 | 850.4 | 70.6 | 1093 | 1 | AL561169 | AL561169 | |
| 15 | 848 | 70.4 | 1099 | 1 | AL582444 | AL582444 | |
| 16 | 844.4 | 70.1 | 1081 | 5 | CR343983 | CR343983 | |
| 17 | 840.6 | 69.8 | 1079 | 5 | EX405259 | EX405259 | |
| 18 | 785 | 65.2 | 971 | 1 | AL570078 | AL570078 | |
| 19 | 783.8 | 65.0 | 933 | 1 | AL543690 | AL543690 | |
| 20 | 763.6 | 63.4 | 920 | 1 | AL539757 | AL539757 | |
| 21 | 763.6 | 63.4 | 944 | 1 | AL582574 | AL582574 | |
| 22 | 747.6 | 62.0 | 1088 | 5 | CR343984 | CR343984 | |
| 23 | 738.8 | 61.3 | 993 | 3 | AL610095 | full-length | |
| 24 | 737.8 | 61.2 | 887 | 1 | AL553980 | AL553980 | |

```
Qy 61 GCTTACCTACCTCCCTCTGCAGGAACCGCGATATGGCTGCGCTGTGCCCCCGCGCG 120
Db 78 GCTTACCTACCTCCCTCTGCAGGAACCGCGATATGGCTGCGCTGTGCCCCCGCGCG 137

Qy 121 ATTTCTCTCCCGCTGCTT-CCCTTCTCTCGGCTTCTGCTCTCTCTCCGCTCCGCATGG 179
Db 138 ATTTCTCTCCCGCTGCTTCCCTTCTCTCGGCTTCTGCTCTCTCTCCGCTCCGCATGG 197

Qy 180 CGGACGGGCTGCACACAA-GGGGCCCTTCCCTGATAGGTCATCTTCTACAAGGT 238
Db 198 CGGACGGGCTGCACACAAAGGGCGCTTCCCTGGATACGGTCATCTTCTACAAGGT 257

Qy 239 CATTTCCAAAAGCAAGTTCGTC-TGGTGAAGTTCGACACCCAGTACCCTACGGTGAGAA 297
Db 258 CATTTCCAAAAGCAAGTTCGTC-TGGTGAAGTTCGACACCCAGTACCCTACGGTGAGAA 317

Qy 298 CGAGATGAGTTCAAGCGTCTT-CTGAAAACTCGGCTTCCAGCGATGATCTTCTGGTGGC 356
Db 318 CGAGATGAGTTCAAGCGTCTTCTGCTGAAAACTCGGCTTCCAGCGATGATCTTCTGGTGGC 377

Qy 357 AGAGTGGGATCTCAGATTAT-GTGACAACTGACATGAGCTGAGTGAGAAATACAA 415
Db 378 AGAGTGGGATCTCAGATTATGAGTGAAGCTGAAGCTGAGTGAGAAATACAA 437

Qy 416 GCTGACAAAGAGAGCTACCCA-TCTTCTACTCTTCCGGATGGGACTTTGAGAACCC 474
Db 438 GCTGACAAAGAGAGCTACCCAGCTTCTACTCTTCCGGATGGGACTTTGAGAACCC 497

Qy 475 AGTCCCATACACTGGGGCAGTT-AGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGG 533
Db 498 AGTCCCATACACTGGGGCAGTTAAGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGG 557

Qy 534 GGTCTACATGATGCTCTGG-TGCCTGCTGTATACAGCCCTTGGCCGGGGAGTTTCA 592
Db 558 GGTCTACATGATGCTCTGGTGGCTGCTGTATACAGCCCTTGGCCGGGGAGTTTCA 617

Qy 593 CAGGCTCTGCTGTGGAGG-CCGCGAGGCTCTTTGAAGCAGGGGCAAGATAACTCTC 651
Db 618 CAGGCTCTGCTGTGGAGGCGCGCAGGCTCTTTGAAGCAGGGGCAAGATAACTCTC 677

Qy 652 AAGTGTGAAGAGACTCAAGAG-AGTGGGCGAGCAATACCTGAAGATCATGGGGAAGAT 710
Db 678 AAGTGTGAAGAGACTCAAGAGAGTGGGCGAGCAATACCTGAAGATCATGGGGAAGAT 737

Qy 711 CTTAGACCAAGGAGGACATT-CCAGCATCAGAGTACAGCATCGCCAGGCTGATTGA 769
Db 738 CTTAGACCAAGGAGGAGCTTCCAGCATCAGAGTACAGCATCGCCAGGCTGATTGA 797

Qy 770 GAAGAAAGATGAGTGAC-GGCAGAAGGAGAGCTCCAGAAGAGCTTTAAACATCCTGAC 828
Db 798 GAAGAAAGATGAGTGACGGGAAGAGAGAGCTCCAGAAGAGCTTTAAACATCCTGAC 857

Qy 829 TGCCTTCAGAAAGAGGGGCGGAGAAAGAGAGAGCTGTAAAAGGCTGTCTGTGATTTTC 888
Db 858 TGCCTTCAGAAAGAGGGGCGGAGAAAGAGAGAGCTGTAAAAGGCTGTCTGTGATTTTC 917

Qy 889 CAGGTTTGGTGGGGTAGGAGGGGAGAGTTAACTCTGCTGTGANTCCCTTGTGGA 948
Db 918 CAGGTTTGGTGGGGTAGGAGGGGAGAGTTAACTCTGCTGTGAGTCCCTTGTGGA 977

Qy 949 ATATAAGGGGYSKGGGAAAGWGGTACTTAACCCACGATTTCTGAGCCCTGAGTATGCT 1008
Db 978 ATATAAGGGGAGTAGTGGGAAAGTGTACTTAACCCACGATTTCTGAGCCCTGAGTATGCT 1037

Qy 1009 GGACATTGATCAATGACCATGCTTTGGATGTCTCTAGCTGTCTGGGATAGCTGG 1068
Db 1038 GGACATTGATCAATGACCATGCTTTGGGATGTCTCTAGCTGTCTGGGATAGCTGG 1097

Qy 1069 AGCACTTACTCAGGTGGCTGGTGAATGACACCTCAGAGGAAATGAGTGTCTATAGAGG 1128
Db 1098 AGCACTTACTCAGGTGGCTGGTGAATGACACCTCAGAGGAAATGAGTGTCTATAGAGG 1157
```

```
Qy 1129 AGAGGAGGTGTTACTGCCAGGTCTTTGACAGATGTAATTCTCAT 1174
Db 1158 AGAGGAGGTGTTACTGCCAGGTCTTTGACAGATGTAATTCTCAT 1203

RESULT 2
CR600542 1152 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CSODI006YCL2 of Placentia Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR600542
VERSION CR600542.1 GI:50481349
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1152)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Redwood City, CA 94061
2 (bases 1 to 1152)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..1152
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI006YCL2"
/tissue_type="Placentia Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 83.1%; Score 1000.8; DB 3; Length 1152;
Best Local Similarity 98.2%; Pred. No. 1.2e-270;
Matches 1131; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

Qy 25 TCAGTGACCGCTGACTCGGGCGGTTCTCCACTATCGTTACCTACCTCCCTCTGCAGGA 84
Db 1 TCAGTGACCGCTGACTCGGGCGGTTCTCCACTATCGTTACCTACCTCCCTCTGCAGGA 60

Qy 85 ACCCGGGATATGCTGCGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTT-CCCT 143
Db 61 ACCCGGGATATGCTGCGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTTCCCT 120

Qy 144 TCTCTGGGCTTCCCTGCTCTCGCTCCGCTCCGCTGGGCGGCGCTGCACACCA-GG 202
Db 121 TCTCTGGGCTTCCCTGCTCTCGCTCCGCTCCGCTGGGCGGCGCTGCACACCAAGG 180

Qy 203 CGCCCTTCCCTGGATACGGTCACCTTTCTACAAGTCATTCCCAAAAGCAAGTTCGTC-T 261
Db 181 CGCCCTTCCCTGGATACGGTCACCTTTCTACAAGTCATTCCCAAAAGCAAGTTCGTC 240

Qy 262 GGTGAAGTTCGACACCCAGTACCCCTACGGTGAAGAGCAGGATGAGTTCAAGCGCTT-C 320
Db 241 GGTGAAGTTCGACACCCAGTACCCCTACGGTGAAGAGCAGGATGAGTTCAAGCGCTTGC 300

Qy 321 TGAAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGGATCTCAGATTAT-G 379
Db 301 TGAAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGGATCTCAGATTATGG 360
```

```
QY 380 TCACAAGCTGAACATCTGAGTGGAGTAAATACAAAGCTGGACAGAGAGCTACCCCA-T 438
Db 361 TCACAAGCTGAACATCTGAGTGGAGTAAATACAAAGCTGGACAGAGAGCTACCCAGT 420
QY 439 CTTTACCTCTTCCGGATGGGACTTTTGAGAACCCAGTCCCATACACTCTGGGCGAGTT-A 497
Db 421 CTTTACCTCTTCCGGATGGGACTTTTGAGAACCCAGTCCCATACACTCTGGGCGAGTTAA 480
QY 498 GGTGGAGCCATCAGCGCTGCTGAAGGGGAAGGGTCTACCTAGTATGCTGG-TG 556
Db 481 GGTGGAGCCATCAGCGCTGCTGAAGGGGAAGGGTCTACCTAGTATGCTGGTGTG 540
QY 557 CTTGCTGTATACGAGCCCTGCGCGGAGTTTATCAGGGGCTCTGTGTGGAGG-CG 615
Db 541 CTTGCTGTATACGAGCCCTGCGCGGAGTTTATCAGGGGCTCTGTGTGGAGGCCG 600
QY 616 CCAGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAG-A 674
Db 601 CCAGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAGAA 660
QY 675 GTGGGCGAGCAATACCTGAAGATCATGGGGAAGATCTTTAGACCAAGGGGAGCACTT-CC 733
Db 661 GTGGGCGAGCAATACCTGAAGATCATGGGGAAGATCTTTAGACCAAGGGGAGCACTTCCC 720
QY 734 AGCATCAGAGATGACACGGATCGCCAGCTGATTGAGAGACAAAGATCAGTGAC-GGCA 792
Db 721 AGCATCAGAGATGACACGGATCGCCAGCTGATTGAGAGACAAAGATGAGTGACGGGAA 780
QY 793 GAAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGCCGA 852
Db 781 GAAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGCCGA 840
QY 853 GAAGGAGGAGCTGTAAGAGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGG 912
Db 841 GAAGGAGGAGCTGTAAGAGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGG 900
QY 913 GGAGAGTTAACTGCTGCTGTGANTCCTTGTGGAATATAGGGGGYMSKGGGAAGW 972
Db 901 GGAGAGTTAACTGCTGCTGTGANTCCTTGTGGAATATAGGGGGTAGTGGGGAAGT 960
QY 973 GGTACTAACCCACGATTCAGACCTCTGAGTATGCTGACATTTGATGCTAACATGACCAT 1032
Db 961 GGTACTAACCCACGATTCAGACCTCTGAGTATGCTGACATTTGATGCTAACATGACCAT 1020
QY 1033 GCTTGGAGTGTCTAGTGTGTCTGGGATAGCTGGAGCACTTACTAGGTGGCTGGTGA 1092
Db 1021 GCTTGGAGTGTCTAGTGTGTCTGGGATAGCTGGAGCACTTACTAGGTGGCTGGTGA 1080
QY 1093 AATGACACTCAGAGGATGATGCTATAGAGAGAGAGAGGATGTACTGCCAGGTC 1152
Db 1081 AATGACACTCAGAGGATGATGCTATAGAGAGAGAGAGGATGTACTGCCAGGTC 1140
QY 1153 TTTGACAGATGT 1164
Db 1141 TTTGACAGATGT 1152
```

```
RESULT 3
LOCUS CR623789
DEFINITION full-length cDNA clone CS01079YL05 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR623789
VERSION CR623789.1. GI:50504596
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1375)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
```

REMARK

Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue
Genoscope.
2 (bases 1 to 1375)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

source

1..1375
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01079YL05"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 82.7%; Score 996.8; DB 3; Length 1375;

Best Local Similarity 98.2%; Pred. No. 1.7e-269;

Matches 1127; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

```
QY 70 CTCCTCTCTGAGGAACCCGGCGATATGGCTGCGCTGTCGCCCGCGCGCATTTCTCTC 129
Db 1 CCTCCCTCTGAGGNAACCCGGCGATATGGCTGCGCTGTCGCCCGCGCGCATTTCTCTC 60
QY 130 CCCGCTGCTT-CCCTTCTCTGGGCTTCTCTCTCTCTCTCTCCGCTCCGCAATGGCGGAGCG 188
Db 61 CCCGCTGCTTCCCTTCTCTGGGCTTCTCTCTCTCTCTCTCCGCTCCGCAATGGCGGAGCG 120
QY 189 CTGCAACACAA-GGCGCCCTTCCCTGGATACGGTCACTTTCTACAGGTCAATCCCA 247
Db 121 CTGCAACACCAAGGGCGCCCTTCCCTGGATACGGTCACTTTCTACAGGTCAATCCCA 180
QY 248 AAGCAAGTTCTCT-TGGTGAAGTTTCGACACCCAGTACCCTACGGTGAAGACAGATGA 306
Db 181 AAGCAAGTTCTCTTGGTGAAGTTTCGACACCCAGTACCCTACGGTGAAGACAGATGA 240
QY 307 GTTCAAGCGTCTT-CTGAAAACCTCGGTTTCAGCGATGATCTTTGGTGGCAGAGTGGG 365
Db 241 GTTCAAGCGTCTTGTGAAAACCTCGGTTTCAGCGATGATCTTTGGTGGCAGAGTGGG 300
QY 366 GATCTCAGATTAT-GTGCAAGCTGAACATGGAGCTGAGTGAGAAATACAGCTGGACAA 424
Db 301 GATCTCAGATTATGGTGAACATGGAGCTGAGTGAGAAATACAGCTGGACAA 360
QY 425 AGAGAGCTACCCA-TCTTCTACCTCTTCCGGATGGGACTTTTGAGAACCCAGTCCCATA 483
Db 361 AGAGAGCTACCCAGTCTTCTACCTCTTCCGGATGGGACTTTTGAGAACCCAGTCCCATA 420
QY 484 CACTGGGCGAGTT-AGGTTGGAGCCATCCAGCGCTGGCTGAAGGGCGCAAGGGGTCTACCT 542
Db 421 CACTGGGCGAGTTAAGGTTGGAGCCATCCAGCGCTGGCTGAAGGGCGCAAGGGGTCTACCT 480
QY 543 AGGTATGCTGG-TGCCTGCTGTATACGACGCCCTTGGCCGGGGAGTTCATCAGGGCCCTC 601
Db 481 AGGTATGCTGGTTCCTGTATACGACGCCCTTGGCCGGGGAGTTCATCAGGGCCCTC 540
QY 602 TGGTGTGGAGG-CGCGCCAGGCCCTTCTGAAGCAGGGGCAAGATACCTCTCAAGTGTGAA 660
Db 541 TGGTGTGGAGGCCCGCCAGGCCCTTCTGAAGCAGGGGCAAGATACCTCTCAAGTGTGAA 600
QY 661 GGAGACTCAGAAG-AGTGGGCGCGCAATACCTGAAAGTATCTGGGGAAGATCTTTAGACCA 719
Db 601 GGAGACTCAGAAGAGTGGGCGCGCAATACCTGAAAGTATCTGGGGAAGATCTTTAGACCA 660
QY 720 AGGGAGCAGCTT-CCAGCATCAGAGATGACAGGGATGCCAGGCTGTGATTGAGAAAGCAA 778
```

Db 661 AGGGAGGACTTCCAGCATCAGAGATGACACGGATCGCCAGGCTGATTGAGAAGAA 720
Qy 779 GATGAGTCAC-GGCAGAGAGGAGCTCCAGAGAGCTTAAACATCCTGAGCTGCTTCCA 837
Db 721 GATGAGTCACGGGAAGAAGGAGGAGCTCCAGAGAGCTTAAACATCCTGAGCTGCTTCCA 780
Qy 838 GAAAGAGGGGCGCAGAAAGAGGAGCTGTGTAAGAGGCTGTCTGTGATTTTCCAGGGTTTG 897
Db 781 GAGAGAGGGGCGCAGAAAGAGGAGCTGTGTAAGAGGCTGTCTGTGATTTTCCAGGGTTTG 840
Qy 898 GTGGGGGTAGGAGAGGGANAGTTAACTGCTGGCTGTGANTCCCTTGTGGAAATATAAGGG 957
Db 841 GTGGGGGTAGGAGAGGGAGTTAACTGCTGGCTGTGAGTCCCTTGTGGAATATAAGGG 900
Qy 958 GGYMSKGGGAAAGWGGTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGGACATTGA 1017
Db 901 GGTAGTGGGAAAGTGGTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGGACATTGA 960
Qy 1018 TGCTAAATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGGATAGCTGGAGCACTTAC 1077
Db 961 TGCTAAATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGGATAGCTGGAGCACTTAC 1020
Qy 1078 TCAGTGTGCTGGTGAATGACACCTCAGAGGAATGAGTGTATAGAGAGAGAGGAG 1137
Db 1021 TCAGTGTGCTGGTGAATGACACCTCAGAGGAATGAGTGTATAGAGAGAGAGGAG 1080
Qy 1138 TGTACTGCCAGGCTTTTGACAGATGTAATTTCTCAATCAATTAAGTTTTCAGTGTTTGG 1197
Db 1081 TGTACTGCCAGGCTTTTGACAGATGTAATTTCTCAATCAATTAAGTTTTCAGTGTTTGG 1140
Qy 1198 TTAAGTGG 1205
Db 1141 TTAAGTGG 1148

CR617833 1139 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DL010YA17 of B cells (Ramos cell line)
Cot 25-normalized of Homo sapiens (human).
CR617833
CR617833.1 GI:50498640
HTC; CNSLT cDNA.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1139)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .1139
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL010YA17"
/tissue_type="B cells (Ramos cell line) Cot 25-normalized"

ORIGIN
Query Match 82.0%; Score 987.8; DB 3; Length 1139;
Best Local Similarity 98.2%; Pred No. 5.7e-267;
Matches 1118; Conservative 5; Mismatches 4; Indels 12; Gaps 12;
/plasmid="pCMVSPORT_6"

Qy 75 CTCTGCAGGAACCCGGCGATATGGCTCCCGCTGTGCCCCCGCGCATTTCTCTCCCGC 134
Db 1 CTCTGCAGGAACCCGGCGATATGGCTCCCGCTGTGCCCCCGCGCATTTCTCTCCCGC 60
Qy 135 TGCTT-CCCTTCTCTCTGGGCTTCTCTCTCTCCGCTCCGATGGCGGAGCGGCTGC 193
Db 61 TGCTTCCCTTCTCTGGGCTTCTCTCTCTCCGCTCCGATGGCGGAGCGGCTGC 120
Qy 194 ACACCAA-GGCGCCCTTCCCTGGATACGGTCACTTCTCAAGGTCAATTCCTCCAAAAGCA 252
Db 121 ACACCAAAGGGCGCTTCCCTGGATACGGTCACTTCTCAAGGTCAATTCCTCCAAAAGCA 180
Qy 253 AGTTCGTC-TGGTGAAGTTGACACCCAGTACCCCTACGGTGAGAAGCAGGATGAGTTCA 311
Db 181 AGTTCGTC-TGGTGAAGTTGACACCCAGTACCCCTACGGTGAGAAGCAGGATGAGTTCA 240
Qy 312 AGCTCTTT-CTGAAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCGAGAGTGGGATCT 370
Db 241 AGCTCTTTGCTGAAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCGAGAGTGGGATCT 300
Qy 371 CAGATTAT-GTGCAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGACAAAAGCA 429
Db 301 CAGATTATGTTGACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGACAAAAGCA 360
Qy 430 GCTACCCA-TCTTCTACCTCTTCCGGGATGGGACTTTTGAGAACCCAGTCCCATACACTG 488
Db 361 GCTACCCAGTCTTCTACCTCTTCCGGGATGGGACTTTTGAGAACCCAGTCCCATACACTG 420
Qy 489 GGGCAGTT-AGTTTGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACCTAGGTA 547
Db 421 GGGCAGTTAAAGTTTGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACCTAGGTA 480
Qy 548 TGCTCTGG-TGCTCCCTGTATACGACCCCTGGCGGGAGTTTCATCAGGCGCTCTGGTG 606
Db 481 TGCTCTGGTGGCTGCTGTATACGACCCCTGGCGGGAGTTTCATCAGGCGCTCTGGTG 540
Qy 607 TGGAGG-CCGCCAGGCGCTTCTGAAGCAGGGCGCAAGATAACCTCTCAAGTGTGAAGGAGA 665
Db 541 TGGAGGCGCGCCAGCGCTTCTGAAGCAGGGCGCAAGATAACCTCTCAAGTGTGAAGGAGA 600
Qy 666 CTCAGAAG-AGTGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGG 724
Db 601 CTCAGAAGAGTGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGG 660
Qy 725 AGCACTT-CCAGCATCAGAGATGACGCGATCGCCAGGCTGATTGAGAAGAACAGATGA 783
Db 661 AGGACTTCCCAGCATCAGAGATGACGCGATCGCCAGGCTGATTGAGAAGAACAGATGA 720
Qy 784 GTGAC-GGCAGAAGGAGGAGCTCCAGAAGAGCTTAAACATCTGACTGCCCTTCCAGAGA 842
Db 721 GTGACGGGAAGAGGAGGAGCTCCAGAAGAGCTTAAACATCTGACTGCCCTTCCAGAGA 780
Qy 843 AGGGGGCGAGAAAGAGGAGCTGTAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGG 902
Db 781 AGGGGGCGAGAAAGAGGAGCTGTAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGG 840
Qy 903 GGTAGGAGGGGAGAGTTAACTCTGCTGCTGANTCCCTTGTGGAATATAAGGGGGYMS 962
Db 841 GGTAGGAGGGGAGAGTTAACTCTGCTGCTGAGTCCCTTGTGGAATATAAGGGGGTAG 900
Qy 963 KGGGAAAAAGWGGTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGGACATTGATGTA 1022
Db 901 TGGGAAAAAGTGTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGGACATTGATGTA 960
Qy 1023 ACATGACCATCTTGGGATGCTCTAGCTGGTCTGGGGATAGCTGGAGCACTTACTCAGG 1082

```
Db 961 ACATGACCATGCTTGGGATGTCCTAGCTGGTCTGGGATAGCTGGAGCACTTACTCAGG 1020
Qy 1083 TGGCTGGTGAATGACACCTCAGAGGATGAGTCTATAGAGAGAGAGAGAGTGTAC 1142
Db 1021 TGGCTGGTGAATGACACCTCAGAGGATGAGTCTATAGAGAGAGAGAGAGTGTAC 1080
Qy 1143 TGCCCGAGTCTTTGACAGATGTAATTCATTCAATTAAGTTTCAGTGTGTTGTTAA 1201
Db 1081 TGCCCGAGTCTTTGACAGATGTAATTCATTCAATTAAGTTTCAGTGTGTTGTTAA 1139

RESULT 5
LOCUS CR602804 1119 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DC009YD12 of Neuroblastoma Cot
25-normalized of Homo sapiens (human).
ACCESSION CR602804
VERSION CR602804.1 GI:50483611
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1119)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1119)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and SmaI sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1119
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC009YD12"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 80.3%; Score 967.8; DB 3; Length 1119;
Best Local Similarity 98.1%; Pred. No. 2.5e-261;
Matches 1098; Conservative 5; Mismatches 4; Indels 12; Gaps 12;
Qy 88 CGGCATATGCTGCGCTGTGCGCGCGCGCGCATTTCTCCCGCTGCTT-CCCTTCT 146
Db 1 CGCGCATATGCTGCGCTGTGCGCGCGCGCATTTCTCCCGCTGCTTCCCTTCT 60
Qy 147 CTGGGCTTCTGCTCTCTCGCTCCGATGCGGCGCGCGCTGACACCAA-GGCGC 205
Db 61 CTGGGCTTCTGCTCTCTCGCTCCGATGCGGCGCGCGCTGACACCAAAGGGCGC 120
Qy 206 CTTCCCTCGATACGGTCACTTTCTACAAGGTCAATCCCAAAAGCAAGTTGCTC-TGGT 264
Db 121 CTTCCCTCGATACGGTCACTTTCTACAAGGTCAATCCCAAAAGCAAGTTGCTTGT 180
Qy 265 GAAGTTCGACACCCAGTACCCCTACGTTGAGAGAGAGAGAGTGTTCAGCGCTT-CTGA 323
Db 181 GAAATTTCGACACCCAGTACCCCTACGTTGAGAGAGAGAGTGTTCAGCGCTTGTGCTGA 240
Qy 324 AAATCGGCTTCCAGGATGATCTCTGTTGGCAGAGTGGGATCTCAGATTAT-GTGA 382
|||||
```

```
241 AAATCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTATGTTGA 300
Qy 393 CAAGCTGAACATGAGAGCTGAGTGAATAACAAGCTGGAACAAGAGAGCTTACCCA-TCCT 441
Db 301 CAAGCTGAACATGAGAGCTGAGTGAATAACAAGCTGGAACAAGAGAGCTTACCCAAGTCTT 360
Qy 442 CTACCTCTTCGGGATGGGACTTTTGAAACCCAGTCCCATACACTGCGGGCAGTT-AGT 500
Db 361 CTACCTCTTCGGGATGGGACTTTTGAAACCCAGTCCCATACACTGCGGGCAGTTAAGT 420
Qy 501 TGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGCTCTACCTAGGTATGCTCTGG-TGCCT 559
Db 421 TGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGCTCTACCTAGGTATGCTCTGGTTCCT 480
Qy 560 GCCTGTATACAGCGCTTGGCGGGGAGTTCATACAGGGCTCTGTTGGTGGAGG-CCGCA 618
Db 481 GCCTGTATACAGCGCTTGGCGGGGAGTTCATCAGGGCTCTGTTGGTGGAGGCGCCGCA 540
Qy 619 GGCCTCTTTGAAGCAGGGGCAAGATAAATCTCTCAAGTGTGAAGGAGACTCAGAAG-AGTG 677
Db 541 GGCCTCTTTGAAGCAGGGGCAAGATAAATCTCTCAAGTGTGAAGGAGACTCAGAAGAGTG 600
Qy 678 GCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGC 736
Db 601 GCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTTCCAGC 660
Qy 737 ATCAGAGATGACAGGATCGCAGCTGATTTGAGAGAAACAAGATGAGTGAC-GCAGAA 795
Db 661 ATCAGAGATGACAGGATCGCAGCTGATTTGAGAGAAACAAGATGAGTGAGTACGGGAAGA 720
Qy 796 GGAGGAGCTCCAGAGAGAGCTTAAACATCTCAGCTGCCTTCCAGAGAGAGGGGGCCGAGAA 855
Db 721 GGAGGAGCTCCAGAGAGAGCTTAAACATCTCAGCTGCCTTCCAGAGAGAGGGGGCCGAGAA 780
Qy 856 AGAGGAGCTGTAAAAAGGCTCTCTGTGATTTTCCAGGGTTTGGTGGGGTATAGGAGGGGA 915
Db 781 AGAGGAGCTGTAAAAAGGCTCTCTGTGATTTTCCAGGGTTTGGTGGGGTATAGGAGGGGA 840
Qy 916 NAGTTAACTCTGCTGCTGTGANTCCCTTGTGGAATATTAAGGGGGMKGGGAAAGWGT 975
Db 841 GAGTTAACTCTGCTGCTGTGANTCCCTTGTGGAATATTAAGGGGGMKGGGAAAGWGT 900
Qy 976 ACTAACCCAGATCTGAGCCCTGAGTATGCTTGACATTCATGCTACATGACCATGCT 1035
Db 901 ACTAACCCAGATCTGAGCCCTGAGTATGCTTGACATTCATGCTACATGACCATGCT 960
Qy 1036 TGGGATGCTCTAGCTGCTCTGGGATAGCTTGGAGCACTTACTCAGGTGGCTGGTGAAT 1095
Db 961 TGGGATGCTCTAGCTGCTCTGGGATAGCTTGGAGCACTTACTCAGGTGGCTGGTGAAT 1020
Qy 1096 GACACCTCAGAGGAATGAGTGTATAGAGAGAGAGAGAGTGTACTGCGCCAGGTCTTT 1155
Db 1021 GACACCTCAGAGGAATGAGTGTATAGAGAGAGAGAGAGTGTACTGCGCCAGGTCTTT 1080
Qy 1156 GACAGATGATTTCTCATTCAATTAAGTTTCAGTGT 1194
Db 1081 GACAGATGATTTCTCATTCAATTAAGTTTCAGTGT 1119

RESULT 6
LOCUS CR611287 1119 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DI072YH08 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR611287
VERSION CR611287.1 GI:50492094
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1119)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
REFERENCE
AUTHORS
```

```
TITLE
JOURNAL
REMARK

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
source
1. .1111
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI072XH08"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match      80.3%; Score 967.8; DB 3; Length 1119;
Best Local Similarity 98.1%; Pred. No. 2.5e-261;
Matches 1098; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

QY 70 CCTCCTCTGCAGGAACCGCGATATGGCTGCGCGTGTGCGCGCGCGCGCATTTCTCTC 129
DB 1 CCTCCTCTGCAGGAACCGCGATATGGCTGCGCGTGTGCGCGCGCGCATTTCTCTC 60

QY 130 CCGCTGCTT-CCCTTCTCTGGGCTTCTGCTCTCTCCGTCGCGATGCGGCGAGCGG 188
DB 61 CCGCTGCTTCCCTTCTCTGGGCTTCTGCTCTCTCCGTCGCGATGCGGCGAGCGG 120

QY 189 CTGTGACACCAA-GGCGCCCTTCCCTCGATACGGTCACTTTCTACAAGGTCATTTCCCAA 247
DB 121 CTGTGACACCAAAGGCGCCCTTCCCTCGATACGGTCACTTTCTACAAGGTCATTTCCCAA 180

QY 248 AAGCAAGTTCGTC-TGGTGAAGTTCGACACCCAGTACCCTTACGGTGAGAACGAGGATGA 306
DB 181 AAGCAAGTTCGTC-TGGTGAAGTTCGACACCCAGTACCCTTACGGTGAGAACGAGGATGA 240

QY 307 GTTCAAGGCTCTT-CTGAAAAGTTCGCGATGATCTCTTGGTGGGAGAGGTTGGG 365
DB 241 GTTCAAGGCTCTTCTGAAAAGTTCGCGATGATCTCTTGGTGGGAGAGGTTGGG 300

QY 366 GATCTCAGATTAT-CTGACAGCTCAACATGAGCTGAGTGAGAAATCAAGCTGACAA 424
DB 301 GATCTCAGATTATGAGTGAAGCTCAACATGAGCTGAGTGAGAAATCAAGCTGACAA 360

QY 425 AGAGAGCTACCCA-TCTTCTACTCTTCCGGATGGGACTTTGAGAACCCAGTCCCAT 483
DB 361 AGAGAGCTACCCAGTCTTCTACTCTTCCGGATGGGACTTTGAGAACCCAGTCCCAT 420

QY 484 CACTGGGAGGTT-AGGTTGAGCCTACGCGCTGAGAGGCGGAGGAGGTTCTACT 542
DB 421 CACTGGGAGGTTAAGGTTGAGCCTACGCGCTGAGAGGCGGAGGAGGTTCTACT 480

QY 543 AGGTATGCTTG-TCGCTGCTGTATACGACGCTGCGCGGAGGTTTCATCAGGCGCTC 601
DB 481 AGGTATGCTTGTTGCTGCTGTATACGACGCTGCGCGGAGGTTTCATCAGGCGCTC 540

QY 602 TGGTGTGGAGG-CCGCCAGGCGCTTCTGAAGCAGGCGCAAGATACCTCTCAAGTGTGAA 660
DB 541 TGGTGTGGAGGCGGCCAGGCGCTTCTGAAGCAGGCGCAAGATACCTCTCAAGTGTGAA 600

QY 661 GGAGACTCAGAAAG-AGTGGGCGGAGCAATACCTGAAGATCATGCGGAGAGATCTTAGACA 719
DB 601 GGAGACTCAGAAAGAGTGGGCGGAGCAATACCTGAAGATCATGCGGAGAGATCTTAGACA 660

Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1119)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .1111
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI043YB01"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_5"

ORIGIN

720 AGGGAGACACTT-CCAGCATCAGAGATGACACCGATCGCCAGGCTGATTGAGAAGACAA 778
DB 661 AGGGAGACACTTCCCCAGCATCAGAGATGACACCGATCGCCAGGCTGATTGAGAAGACAA 720

QY 779 GATGAGTGAC-GGCAGAAAGAGGAGGCTCCAGAGAGCTTAAACATCTCTGACTGCTTCCA 837
DB 721 GATGAGTGACGGGGAAGAGGAGGCTCCAGAGAGCTTAAACATCTCTGACTGCTTCCA 780

QY 838 GAAGAAGGGGCCGAGAAAGAGGAGCTGTAAAGAGGCTGTCTGTGATTTTCCAGGGTTTG 897
DB 781 GAAGAAGGGGCCGAGAAAGAGGAGCTGTAAAGAGGCTGTCTGTGATTTTCCAGGGTTTG 840

QY 898 GTGGGGGTAGGGAGGGGAGGAGTTAACTGCTGCTGGCTGTGANTCCCTTGTGAATATAAGG 957
DB 841 GTGGGGGTAGGGAGGGGAGGAGTTAACTGCTGCTGGCTGTGANTCCCTTGTGAATATAAGG 900

QY 958 GGYMSKGGGAAAAGWGGTACTAACCCACGATTCGAGCCCTGAGTATGCCCTGGACATTGA 1017
DB 901 GGTAGTGGGAAAAGTGGTACTTAACCCACGATTCGAGCCCTGAGTATGCCCTGGACATTGA 960

QY 1018 TGCTAACATGACCATGCTTGGGATGCTCTAGCTGCTGGGATAGCTGGGAGCATTAC 1077
DB 961 TGCTAACATGACCATGCTTGGGATGCTCTAGCTGCTGGGATAGCTGGGAGCATTAC 1020

QY 1078 TCAGGTGGCTGGTGAATATGACACCTCAGAAAGGATGAGTGTCTATAGAGAGGAGAGGAG 1137
DB 1021 TCAGGTGGCTGGTGAATATGACACCTCAGAAAGGATGAGTGTCTATAGAGAGGAGAGGAG 1080

QY 1138 TGTACTGCCAGGCTCTTTGACAGATGTAATTTCTCAATCA 1176
DB 1081 TGTACTGCCAGGCTCTTTGACAGATGTAATTTCTCAATCA 1119

RESULT 7
LOCUS
DEFINITION
CR604444 1111 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CSODI043YB01 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION
CR604444
VERSION
CR604444.1 GI:50485251
KEYWORDS
HTC; CNSUT_CDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1111)
AUTHORS
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1111)
Genoscope.
Direct Submission
TITLE
JOURNAL

COMMENT
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. .1111
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI043YB01"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_5"

FEATURES
source
```



```
Query Match 79.7%; Score 959.8; DB 3; Length 1111;
Best Local Similarity 98.1%; Pred. No. 4.5e-259;
Matches 1090; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

Qy 88 CGCGGATATGGCTGCGCTGTCCTCCGCTCCGATGCGGCGAGCGGCTGCACACAA-GGCGC 205
Db 1 CGCGGATATGGCTGCGCTGTCCTCCGCTCCGATGCGGCGAGCGGCTGCACACAAAGGGGC 120
Qy 147 CTGGGCTTCCTGCTCTCTCCGCTCCGATGCGGCGAGCGGCTGCACACAA-GGCGC 205
Db 61 CTGGGCTTCCTGCTCTCTCCGCTCCGATGCGGCGAGCGGCTGCACACAAAGGGGC 120
Qy 206 CTTCCCTCCGATGAGTACGCTGCTTCTCAAGAGTCAATCCCAAAAGCAAGTTCGTC-TGGT 264
Db 121 CTTCCCTCCGATGAGTACGCTGCTTCTCAAGAGTCAATCCCAAAAGCAAGTTCGTC-TGGT 180
Qy 265 GAAGTTCCGACACCCAGTACCCCTCCGCTCCGATGCGGCGAGCGGCTGCACACAA-GGCGC 323
Db 181 GAAGTTCCGACACCCAGTACCCCTCCGCTCCGATGCGGCGAGCGGCTGCACACAA-GGCGC 240
Qy 324 AAATCGGCTTCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTAT-GTGA 382
Db 241 AAATCGGCTTCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTAT-GTGA 300
Qy 383 CAAGCTGACATGAGCTGAGTGAGAAATACAGCTGACAGAGAGAGCTACCCA-TCTT 441
Db 301 CAAGCTGACATGAGCTGAGTGAGAAATACAGCTGACAGAGAGAGCTACCCA-TCTT 360
Qy 442 CTACCTCTTCGCGATGGGAGCTTTGAGAACCCAGTCCCATACACTGGGCGAGTT-AGGT 500
Db 361 CTACCTCTTCGCGATGGGAGCTTTGAGAACCCAGTCCCATACACTGGGCGAGTT-AGGT 420
Qy 501 TGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGTCTACCTAGTATGCCCTGG-TGCT 559
Db 421 TGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGTCTACCTAGTATGCCCTGG-TGCT 480
Qy 560 GCCTGTATACAGCGCTGGCGGGAGTTCATCAGGGCTCTGTGTGGAGG-CCGCA 618
Db 481 GCCTGTATACAGCGCTGGCGGGAGTTCATCAGGGCTCTGTGTGGAGG-CCGCA 540
Qy 619 GGCCTCTTTGAAGCAGGGGCAAGATAAAGCTCTCAAGTGTGAAGGAGACTCAGAAG-AGTG 677
Db 541 GGCCTCTTTGAAGCAGGGGCAAGATAAAGCTCTCAAGTGTGAAGGAGACTCAGAAG-AGTG 600
Qy 678 GCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGC 736
Db 601 GCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGC 660
Qy 737 ATCAGAGATGACACCGATGCGCAGGCTGATTTGAGAGAAACAGATGAGTGAC-GGCAGAA 795
Db 661 ATCAGAGATGACACCGATGCGCAGGCTGATTTGAGAGAAACAGATGAGTGAGTGGGAGGAA 720
Qy 796 GGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGGCGAGAA 855
Db 721 GGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGGCGAGAA 780
Qy 856 AGAGGAGCTGAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAGGGGA 915
Db 781 AGAGGAGCTGAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAGGGGA 840
Qy 916 NAGTTAACCTGCTGCTGTGANTCCCTGTGTGAAATATAGGGGGYMSGGGAAAGWGGT 975
Db 841 GAGTTAACCTGCTGCTGTGANTCCCTGTGTGAAATATAGGGGGGTAGTGGGAAAGTGGT 900
Qy 976 ACTACCCACGATCTGAGCCCTGAGTATGCTGACATTTGATGCTTACATGACCATGCT 1035
Db 901 ACTACCCACGATCTGAGCCCTGAGTATGCTGACATTTGATGCTTACATGACCATGCT 960
Qy 1036 TGGGATGCTCTAGCTGGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGAAT 1095
Db 961 TGGGATGCTCTAGCTGGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGAAT 1020
```

```
Qy 1096 GACACCTCAGAGGAATGAGTGTCTATAGAGAGGAGAGGAGTGTACTGCCAGGCTCTTT 1155
Db 1021 GACACCTCAGAGGAATGAGTGTCTATAGAGAGGAGAGGAGTGTACTGCCAGGCTCTTT 1080
Qy 1156 GACAGATGAATTTCTCAATTCATTAAGTTT 1186
Db 1081 GACAGATGAATTTCTCAATTCATTAAGTTT 1111

RESULT 8
CR617687 1111 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CSODL001YK02 of B cells (Ramos cell line)
DEFINITION Cot 25-normalized of Homo sapiens (human).
ACCESSION CR617687.1 GI:50498494
VERSION HTC; CNSLT cDNA.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1111)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1111)
REFERENCE 2 (bases 1 to 1111)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..1111
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODL001YK02"
/tissue type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 79.7%; Score 959.8; DB 3; Length 1111;
Best Local Similarity 98.1%; Pred. No. 4.5e-259;
Matches 1090; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

Qy 88 CGCGGATATGGCTGCGCTGTCCTCCGCTCCGATGCGGCGAGCGGCTGCACACAA-GGCGC 205
Db 1 CGCGGATATGGCTGCGCTGTCCTCCGCTCCGATGCGGCGAGCGGCTGCACACAAAGGGGC 120
Qy 147 CTGGGCTTCCTGCTCTCTCCGCTCCGATGCGGCGAGCGGCTGCACACAA-GGCGC 205
Db 61 CTGGGCTTCCTGCTCTCTCCGCTCCGATGCGGCGAGCGGCTGCACACAAAGGGGC 120
Qy 206 CTTCCCTCCGATGAGTACGCTGCTTCTCAAGAGTCAATCCCAAAAGCAAGTTCGTC-TGGT 264
Db 121 CTTCCCTCCGATGAGTACGCTGCTTCTCAAGAGTCAATCCCAAAAGCAAGTTCGTC-TGGT 180
Qy 265 GAAGTTCCGACACCCAGTACCCCTCCGCTCCGATGCGGCGAGCGGCTGCACACAA-GGCGC 323
Db 181 GAAGTTCCGACACCCAGTACCCCTCCGCTCCGATGCGGCGAGCGGCTGCACACAA-GGCGC 240
Qy 324 AAATCGGCTTCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTAT-GTGA 382
Db 241 AAATCGGCTTCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTAT-GTGA 300
Qy 383 CAAGCTGACATGAGCTGAGTGAGAAATACAGCTGACAGAGAGAGCTACCCA-TCTT 441
Db 301 CAAGCTGACATGAGCTGAGTGAGAAATACAGCTGACAGAGAGAGCTACCCA-TCTT 360
Qy 442 CTACCTCTTCGCGATGGGAGCTTTGAGAACCCAGTCCCATACACTGGGCGAGTT-AGGT 500
Db 361 CTACCTCTTCGCGATGGGAGCTTTGAGAACCCAGTCCCATACACTGGGCGAGTT-AGGT 420
Qy 501 TGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGTCTACCTAGTATGCCCTGG-TGCT 559
Db 421 TGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGTCTACCTAGTATGCCCTGG-TGCT 480
Qy 560 GCCTGTATACAGCGCTGGCGGGAGTTCATCAGGGCTCTGTGTGGAGG-CCGCA 618
Db 481 GCCTGTATACAGCGCTGGCGGGAGTTCATCAGGGCTCTGTGTGGAGG-CCGCA 540
Qy 619 GGCCTCTTTGAAGCAGGGGCAAGATAAAGCTCTCAAGTGTGAAGGAGACTCAGAAG-AGTG 677
Db 541 GGCCTCTTTGAAGCAGGGGCAAGATAAAGCTCTCAAGTGTGAAGGAGACTCAGAAG-AGTG 600
Qy 678 GCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGC 736
Db 601 GCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGC 660
Qy 737 ATCAGAGATGACACCGATGCGCAGGCTGATTTGAGAGAAACAGATGAGTGAC-GGCAGAA 795
Db 661 ATCAGAGATGACACCGATGCGCAGGCTGATTTGAGAGAAACAGATGAGTGAGTGGGAGGAA 720
Qy 796 GGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGGCGAGAA 855
Db 721 GGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGGCGAGAA 780
Qy 856 AGAGGAGCTGAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAGGGGA 915
Db 781 AGAGGAGCTGAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAGGGGA 840
Qy 916 NAGTTAACCTGCTGCTGTGANTCCCTGTGTGAAATATAGGGGGYMSGGGAAAGWGGT 975
Db 841 GAGTTAACCTGCTGCTGTGANTCCCTGTGTGAAATATAGGGGGGTAGTGGGAAAGTGGT 900
Qy 976 ACTACCCACGATCTGAGCCCTGAGTATGCTGACATTTGATGCTTACATGACCATGCT 1035
Db 901 ACTACCCACGATCTGAGCCCTGAGTATGCTGACATTTGATGCTTACATGACCATGCT 960
Qy 1036 TGGGATGCTCTAGCTGGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGAAT 1095
Db 961 TGGGATGCTCTAGCTGGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGAAT 1020
```

QY 383 CAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGCTACCCA-TCTT 441
Db 301 CAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGCTACCCAGTCTT 360
QY 442 CTACCTCTTCGGGATGGGGATTTGAGAACCCAGTCCCATACATACACTGGGGCAGTT-AGGT 500
Db 361 CTACCTCTTCGGGATGGGGATTTGAGAACCCAGTCCCATACATACACTGGGGCAGTTAAGGT 420
QY 501 TGGAGCCATCAGCGCTCGCTGAAGGGGCAAGGGTCTACCTAGTATGCTGG-TGCCT 559
Db 421 TGGAGCCATCAGCGCTCGCTGAAGGGGCAAGGGTCTACCTAGTATGCTGGTGGCT 480
QY 560 GCCTGTATACAGACGCTTCGGCGGGAGTTTCATCAGGGCTCTGTGTGGAGG-CCGCCA 618
Db 481 GCCTGTATACAGACGCTTCGGCGGGAGTTTCATCAGGGCTCTGTGTGGAGGCGGCCA 540
QY 619 GGCCCTCTTGAAGCAGGGGCAAGATACTCTCAAGTGTGAAGGAGCTCAGAAG-AGTG 677
Db 541 GGCCCTCTTGAAGCAGGGGCAAGATACTCTCAAGTGTGAAGGAGCTCAGAAGAGTG 600
QY 678 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGGACTTCCAGC 736
Db 601 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGGACTTCCAGC 660
QY 737 ATCAGAGATGACCGGATCGCCAGCTGANTTGAGAGAAACAAGATGAGTGAC-GGCAGAA 795
Db 661 ATCAGAGATGACCGGATCGCCAGCTGANTTGAGAGAAACAAGATGAGTGAGCGGAGAA 720
QY 796 GGAGAGCTCAGAAAGACTTAAACATCTGACTGCTTCAGAAAGAGGGGGCGGAGAA 855
Db 721 GGAGAGCTCAGAAAGACTTAAACATCTGACTGCTTCAGAAAGAGGGGGCGGAGAA 780
QY 856 AGAGAGCTGTAAAAGGCTGTCTGATTTTCAGGGTTTGGTGGGGTAGGGAGGGGA 915
Db 781 AGAGAGCTGTAAAAGGCTGTCTGATTTTCAGGGTTTGGTGGGGTAGGGAGGGGA 840
QY 916 NAGTTAACTGTGCTGTGANTCCCTTTGTGGAATATAAGGGGGMKSGGAAAGMGST 975
Db 841 GAGTTAACTGTGCTGTGANTCCCTTTGTGGAATATAAGGGGTTAGTGGAAAGTGGT 900
QY 976 ACTAACCACAGATTCAGACCTCAGTATGCTGACATTTGATGCTAATCAGTACCATGCT 1035
Db 901 ACTAACCACAGATTCAGACCTCAGTATGCTGACATTTGATGCTAATCAGTACCATGCT 960
QY 1036 TGGGATGCTCTAGCTGTCTGGGATAGCTGGACACTTACTCAGTGGCTGGTGAAT 1095
Db 961 TGGGATGCTCTAGCTGTCTGGGATAGCTGGACACTTACTCAGTGGCTGGTGAAT 1020
QY 1096 GACACCTCAGAGGAATGAGTCTATAGAGAGGAGAGAGTGTPACTGCCAGTCTTTT 1155
Db 1021 GACACCTCAGAGGAATGAGTCTATAGAGAGGAGAGAGTGTPACTGCCAGTCTTTT 1080
QY 1156 GACAGATGTAATCTCATTTAAAGTTT 1186
Db 1081 GACAGATGTAATCTCATTTAAAGTTT 1111

RESULT 9
CR603906
LOCUS
DEFINITION
full-length cDNA clone CS0DH002YC21 of T cells (Jurkat cell line)
of Homo sapiens (human).
ACCESSION
CR603906
VERSION
CR603906.1 GI:50484713
KEYWORDS
HTC; CNSLT_cDNA.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1110)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished

REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1110)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1110
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH002YC21"
/tissue_type="T cells (Jurkat cell line)"
/plasmid="pCMVSPORT_6"

Query Match 79.6%; Score 958.8; DB 3; Length 1110;
Best Local Similarity 98.1%; Pred. No. 8.6e-259;
Matches 1089; Conservative 5; Mismatches 4; Indels 12; Gaps 12;
QY 88 CGGCGATATGGTCCCGCTGTGCCCGCGCGCATTTCTCCCGCGTTCCTTCCCGCTTCT 146
Db 1 CGGCGATATGGTCCCGCTGTGCCCGCGCGCATTTCTTCTCCCGCTTCTTCTCCCGCTTCT 60
QY 147 CCTGGGCTTCTGCTCTCTCCGCTCCGCGCATGGCGGAGCGGCTGCACACCAA-GGCGC 205
Db 61 CTGGGGTCTCTGCTCTCTCCGCTCCGCGCATGGCGGAGCGGCTGCACACCAAGGGCGC 120
QY 206 CTTTCCCTCGATACGCTGACTTTTCTCAAGGTCAITTCCTCCAAAAGCAAGTTTCGT- 264
Db 121 CTTTCCCTCGATACGCTGACTTTTCTCAAGGTCAITTCCTCCAAAAGCAAGTTTCGTTCGT 180
QY 265 GAAGTTGCACACCCAGTACCCCTACCGTGGAGAGCAGGAGTTCAGCGCTCTT-CTGA 323
Db 181 GAAGTTGCACACCCAGTACCCCTACCGTGGAGAGCAGGAGTTCAGCGCTCTTTCGTGA 240
QY 324 AAATCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTAT-GTGA 382
Db 241 AAATCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTATGTGA 300
QY 383 CAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGGAACAAGAGAGCTACCCA-TCTT 441
Db 301 CAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGGAACAAGAGAGCTACCCAGTCTT 360
QY 442 CTACCTCTTCGGGATGGGGATTTGAGAACCCAGTCCCATACACTGGGGCAGTT-AGGT 500
Db 361 CTACCTCTTCGGGATGGGGATTTGAGAACCCAGTCCCATACACTGGGGCAGTTAAGGT 420
QY 501 TGGAGCCATCAGCGCTCGCTGAAGGGGCAAGGGTCTTACCTAGTATGCTGG-TGCCT 559
Db 421 TGGAGCCATCAGCGCTCGCTGAAGGGGCAAGGGTCTTACCTAGTATGCTGGTGGCT 480
QY 560 GCCTGTATACAGACGCTTCGGCGGGAGTTTCATCAGGGCTCTGTGTGGAGG-CCGCCA 618
Db 481 GCCTGTATACAGACGCTTCGGCGGGAGTTTCATCAGGGCTCTGTGTGGAGGCGGCCA 540
QY 619 GGCCCTCTTGAAGCAGGGGCAAGATACTCTCAAGTGTGAAGGAGCTCAGAAG-AGTG 677
Db 541 GGCCCTCTTGAAGCAGGGGCAAGATACTCTCAAGTGTGAAGGAGCTCAGAAGAGTG 600
QY 678 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGGACTT-CCAGC 736
Db 601 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGGACTTCCAGC 660
QY 737 ATCAGAGATGACCGGATCGCCAGCTGANTTGAGAGAAACAAGATGAGTGAC-GGCAGAA 795

|||||
661 ATCAGAGATGACACGGATCCGACGCTGATTTGAGAGAACACAGATGATGACGGAGAA 720
796 GGAGGAGCTCCAGAGAGCTTAAACATCTGATCCCTTCCAGAGAAAGGGGCGAGAA 855
721 GGAGGAGCTCCAGAGAGCTTAAACATCTGATCCCTTCCAGAGAAAGGGGCGAGAA 780
856 AGAGGAGCTGTAAGAAGCTGCTGATTTTCCAGGGGTTGGTGGGGTAGGGAGGGA 915
781 AGAGGAGCTGTAAGAAGCTGCTGATTTTCCAGGGGTTGGTGGGGTAGGGAGGGA 840
916 NAGTTAACTCTGCTGCTGTGATCCCTTGTGGAATATAAGGGGYYMSKGGGAAAGWGT 975
841 GAGTTAACTCTGCTGCTGTGATCCCTTGTGGAATATAAGGGGTTAGTGGGAAAGTGT 900
976 ACTAACCCAGCATCTGAGCCCTGATGATGCTTGACATTAATGATTAACGACCATGCT 1035
901 ACTAACCCAGCATCTGAGCCCTGATGATGCTTGACATTAATGATTAACGACCATGCT 960
1036 TGGGATGCTCTAGCTGCTGGGATAGCTGAGCATTACTCAGGTGGCTGGTGAAT 1095
961 TGGGATGCTCTAGCTGCTGGGATAGCTGAGCATTACTCAGGTGGCTGGTGAAT 1020
1096 GACACCTCAGAGGAATGAGTCTATAGAGAGGAGAGGAGTGTACTGCCAGGTCTTT 1155
1021 GACACCTCAGAGGAATGAGTCTATAGAGAGGAGAGGAGTGTACTGCCAGGTCTTT 1080
1156 GACAGATGAATTTCTCATTTAAAGTT 1185
1081 GACAGATGAATTTCTCATTTAAAGTT 1110

RESULT 10
BX403042
LOCUS BX403042 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CSODL010YA17 5-PRIME, mRNA sequence.
ACCESSION BX403042
VERSION BX403042.2 GI:46844967
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1137)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30611153.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6716.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna/cna=CSIAL003ZC09QPI&c=6716.f.

FEATURES

source

1. 1137
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODL010YA17"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was

digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 77.0%; Score 928.4; DB 5; Length 1137;
Best Local Similarity 94.8%; Pred. No. 3.3e-250;
Matches 1080; Conservative 23; Mismatches 22; Indels 14; Gaps 14;
QY 75 CTCTGCAGAAACCCGGCGATATGGCTGCCGTGTGCCCCCGCGCCATTTCTCTCCCCGC 134
Db 1 CTCTGCAGAAACCCGGCGATATGGCTGCCGTGTGCCCCCGCGCCATTTCTCTCCCCGC 60
QY 135 TGCTTT-CCCTTCTCTCTGGGCTTCTGCTCTCTCGCTCCGCATCGCGCGCGCGCTGC 193
Db 61 TGCTTCCCTTCTCTCTGGGCTTCTGCTCTCTCGCTCCGCATCGCGCGCGCTGC 120
QY 194 ACACAA-GGCGCCCTTCCCTCGGATACGGTCACTTTCTACAAGGTCAATTCCTCAAAAGCA 252
Db 121 ACACNAGGGCGCTTCCCTCGGATACGGTCACTTTCTACAAGGTCAATTCCTCAAAAGCA 180
QY 253 AGTTCGTC-TGGTGAAGTTTCACACCCAGTACCCTACGGTGAGAAAGCAGGATGATTC 311
Db 181 AGTTCGTC-TGGTGAAGTTTCACACCCAGTACCCTACGGTGAGAAAGCAGGATGATTC 240
QY 312 AGCTCTTT-CTGAAAACCTCGGCTTCCAGCGATGATCTCTTGGTGCAGAGTGGGGATCT 370
Db 241 AGCTCTTTCTGTAAGTTCGGCTTCCAGCGATGATCTCTTGGTGCAGAGTGGGGATCT 300
QY 371 CAGATTAT-GTGACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGCACAAAGAGA 429
Db 301 CAGATTATGTTGACACAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGCACAAAGAGA 360
QY 430 GCTACCCA-TCTTTTACCTTCTCCGGGATGGGACTTTTGAGAACCCAGTCCCATACATG 488
Db 361 GCTACCCAGTCTTCTACCTTCTCCGGGATGGGACTTTTGAGAACCCAGTCCCATACATG 420
QY 489 GGGCAGTT-AGTTGGAGCCATCCAGCGCTGGCTGAGGGGCAAGGGTCTACCTAGTA 547
Db 421 GGGCAGTTAAGTTTGGAGCCATCCAGCGCTGGCTGAGGGGCAAGGGTCTACCTAGTA 480
QY 548 TGCCTGG-TGCCTGCTGTATACGACGCCCTGGCGGGGAGTTTCAATCAGGGCTCTCGTG 606
Db 481 TGCCTGGTTGCTGCGGTATACGACGCCCTGGCGGGGAGTTTCAATCAGGGCTCTCGTG 540
QY 607 TGGAGG-CCGCGAGGCCCTTCTTGAAGCAGGGGCAAGATAAACCCTCTCAAGTGTGAAGAGA 665
Db 541 TGGAGGCCGCGCAGGCCCTTCTTGAAGCAGGGGCAAGATAAACCCTCTCAAGTGTGAAGAGA 600
QY 666 CTCAGAG-AGTGGCGCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGG 724
Db 601 CTCAGAGAGTGGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGG 660
QY 725 AGCACTT-CCAGCATCAGAGATGACAGCGATCGCCAGGCTGATTTGAGAGAACCAAGATGA 783
Db 661 AGCACTTCCAGCATCAGAGATGACAGCGATCGCCAGGCTGATTTGAGAGAACCAAGATGA 720
QY 784 GTGAC-GGCGAGAGGAGAGCTCCAGAGAGCTTTAAACATCTGACTGCCCTTCCAGAGA 842
Db 721 GTGACGGGAAGAGAGGAGCTCCAGAGAGCTTTAAACATCTGACTGCCCTTCCAGAGA 780
QY 843 AGGGGGCGAGAAAGAGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTTGGTGGG 902
Db 781 AGGGGGCGAGAAAGAGAGCTGTAAAAGGCTGTGTGTTATTTTCCAGGGTTGGTGGG 840
QY 903 GGTAGGGAGGAGAGTTAACTCTGCTGTGANTCCCTTGTGGAATATTAAGGGGGYMS 962
Db 841 GGTAGGGAGGAGAGTTAACTCTGCTGTGTTAGTTCTTGTGGAATATTAAGGGGGTAG 900
QY 963 KGGGAAAGWGTACTAAACCCAGCATCTGAGCCCTGAGTATGCTGAGCATTTGATGCTA 1022
Db 901 TGGGAAAGTGTACTAAACCCAGCATCTGAGCCCTGAGTATGCTGAGCATTTGATGCTA 960
QY 1023 ACATGACCATCTTGGGAGTGTCTAGCTGCTGCGGATAGCTGGAGCACTTACTCAGG 1082

| | | | |
|------|----|---|------|
| 961 | Db | AMATGACCATGCTTGGGATGCTCTAGCTGCTGGGATAGCTGGAGCMCTTACTCAGG | 1020 |
| 1083 | Qy | TGGCTGCTGTAATGACACCTCAGAAAGCAATGAGTGCCTATAGAGAGGAGAGAGAGTGTAC | 1142 |
| 1021 | Db | TGGCTGCTGTAATGACACMCTCAGAAAGCAATGAGTGCCTATA-AAAAGRRAAAGATTTC | 1079 |
| 1143 | Qy | TGCCAGGCTCTTTGACAGATGTAAATCTCAATCAATTAAGTTCAGTGTGGTTAA | 1201 |
| 1080 | Db | TGCCACGCTCTTTTACAAAGTWAATTTTMTT-MATTAARTTTCACTKTTTGTGTTAA | 1137 |

RESULT 11
CR602432
LOCUS
DEFINITION
CR602432 1062 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CSODF036XH02 of Fetal brain of Homo sapiens
(human).
ACCESSION
CR602432
VERSION
CR602432.1 GI:50483239
KEYWORDS
HTC; CNSLT cDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1062)
AUTHORS
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

| Query Match | 75.6%; | Score 910.8; | DB 3; | Length 1062; |
|-----------------------|-----------------|--|--|--------------|
| Best Local Similarity | 98.0%; | Pred. No. 3e-245; | | |
| Matches 1041; | Conservative 5; | Mismatches 4; | Indels 12; | Gaps 12; |
| Qy | 88 | CGGGATATGCGTGC | CGCGTGTGCGCCGCGCGCATTTCTCTCCCGGTGCTT | -CCCTTCT 146 |
| Db | 1 | CGGGATATGCGTGC | CGCGTGTGCGCCGCGCATTTCTCTCCCGGTGCTT | CCCTTCT 60 |
| Qy | 147 | CCTGGGCTTCTGCTCT | CTCCGCTTCGCGATGGCGGAGCGGCGTGCACACCAA | -GGGCG 205 |
| Db | 61 | CCTGGGCTTCTGCTCT | CTCCGCTTCGCGATGGCGGAGCGGCGTGCACACCAAGGGCGC | 120 |
| Qy | 206 | CTTTCCCTCGGATACGGTCACTTTCTACAAGGTCAATTCCCAAAAGCAAGTTCGTC | -TGGT 264 | |
| Db | 121 | CTTTCCCTCGGATACGGTCACTTTCTACAAGGTCAATTCCCAAAAGCAAGTTCGTC | TTGGT 180 | |
| Qy | 265 | GAA GTTCGACACCCAGTACCCCTACGGTGTAGAGCAGGATGAGTTCAGCGTCTTT | -CTGA 323 | |
| Db | 181 | GAA GTTCGACACCCAGTACCCCTACGGTGTAGAGCAGGATGAGTTCAGCGTCTTTGTCTGA | 240 | |
| Qy | 324 | AAACTCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTAT | -GTGA 382 | |

| | | | | |
|------|----|--|---|------|
| 241 | Db | | AAACTCGGCTTCAGCGATGATCTCTGTGTGGCAGAGGTGGGGATCTCAGATTATGGTGA | 300 |
| 383 | Qy | | CAAGCTGAACATGGAGCTGAGTCAGAAATAACAAGCTTGGACAAGAGAGAGCTTACCCA-TCCTT | 441 |
| 301 | Db | | CAAGCTGAACATGGAGCTGAGTCAGAAATAACAAGCTTGGACAAGAGAGAGCTTACCCAGTCCT | 360 |
| 442 | Qy | | CTACCTCTTCCGGATGGGGACTTTGAGAACCCAGTCCCATACACTCTGGGGCAGTT-AGGT | 500 |
| 361 | Db | | CTACCTCTTCCGGATGGGGACTTTGAGAACCCAGTCCCATACACTCTGGGGCAGTTAAGGT | 420 |
| 501 | Qy | | TGGAGCCATCCAGCGCTGGCTGAAGGGGCNAAGGGTCTACCTTAGGTATGSCCTGG-TGCCT | 559 |
| 421 | Db | | TGGAGCCATCCAGCGCTGGCTGAAGGGGGCAAGGGGTCTACCTTAGGTATGCTGTGGTGCCT | 480 |
| 560 | Qy | | GCCTGTATACGACGCGCTGGCCGGGAGTTTCATCAGGGCCTCTGTGTGTGGAGG-CGCGCCA | 618 |
| 481 | Db | | GCCTGTATACGACGCGCTGGCCGGGAGTTTCATCAGGGCCTCTGTGTGTGGAGGCCCGCCA | 540 |
| 619 | Qy | | GGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAG-AGTG | 677 |
| 541 | Db | | GGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAGAAGTG | 600 |
| 678 | Qy | | GGCCGAGCAATACCTGAAGATCATTGGGGAAGATCTTTAGACCAAGGGGAGACATT-CCAGC | 736 |
| 601 | Db | | GGCCGAGCAATACCTGAAGATCATTGGGGAAGATCTTTAGACCAAGGGGAGACATTCACAGC | 660 |
| 737 | Qy | | ATCAGAGATACACAGGATCGCCAGGCTGATTTGAGAAGAACAGATGATGTAC-GGCAGAA | 795 |
| 661 | Db | | ATCAGAGATACACAGGATCGCCAGGCTGATTTGAGAAGAACAGATGATGATGACGGGAAGAA | 720 |
| 796 | Qy | | GGAGGAGCTCCAGAAGACTTAAACATCTCTGACTGCCTTCAGAAGAAAGGGGGCCGAGAA | 855 |
| 721 | Db | | GGAGGAGCTCCAGAAGACTTAAACATCTCTGACTGCCTTCAGAAGAAAGGGGGCCGAGAA | 780 |
| 856 | Qy | | AGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGTGTGGGGTAGGGAGGGGA | 915 |
| 781 | Db | | AGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGTGTGGGGTAGGGAGGGGA | 840 |
| 916 | Qy | | NAGTTAACTGTCTGGCTGTGANTCCCTCTGTGGAATATAGGGGGGYMSKGGGAAAAAGWGT | 975 |
| 841 | Db | | GAGTTAACTGTCTGGCTGTGAGTCCCTTGTGGAATATAGGGGGTAGGGGAAAAAGTGGT | 900 |
| 976 | Qy | | ACTAAACCCAGATTCTGAGCCCTGAGTATGCCCTGGACATTGATGCTAAACATGACCATGCT | 1035 |
| 901 | Db | | ACTAAACCCAGATTCTGAGCCCTGAGTATGCCCTGGACATTGATGCTAAACATGACCATGCT | 960 |
| 1036 | Qy | | TGGGATGTCTTAGCTGTCTGGGGTAGCTGGAGCACTTTACTCAGGTGCTGTGTGAAT | 1095 |
| 961 | Db | | TGGGATGTCTTAGCTGTCTGGGGTAGCTGGAGCACTTTACTCAGGTGCTGTGTGAAT | 1020 |
| 1096 | Qy | | GACACCTCAGAAGGAATGAGTGTCTATAGAGAGAGAGAGAG | 1137 |
| 1021 | Db | | GACACCTCAGAAGGAATGAGTGTCTATAGAGAGAGAGAGAG | 1062 |

| | | | | |
|------------|---|-------------|--------|-----------------|
| RESULT 12 | CR594614 | 1058 bp | linear | HTC 21-JUL-2004 |
| LOCUS | CR594614 | | mRNA | |
| DEFINITION | Full-length cDNA clone CS9DJ008YF23 of T cells (Jurkat cell line) Cot 10-normalized of Homo sapiens (human). | | | |
| ACCESSION | CR594614 | | | |
| VERSION | CR594614.1 | GI:50475421 | | |
| KEYWORDS | HTC; CNSLT_CDNA. | | | |
| SOURCE | Homo sapiens (human) | | | |
| ORGANISM | Homo sapiens | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| REFERENCE | 1 (bases 1 to 1058) | | | |
| AUTHORS | Li, W.B., Gruber, C., Jesse, J. and Polayes, D. | | | |
| TITLE | Full-length cDNA libraries and normalization | | | |
| JOURNAL | Unpublished | | | |
| REMARK | Contact : Feng Liang Email : fliang@lifetech.com URL : | | | |

```
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1058)
REFERENCE
Genoscope.
AUTHORS
Direct Submission
TITLE
JOURNAL
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1058
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ008YF23"
/tissue_type="T cells (Jurkat cell line) Cot
10-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 75.3%; Score 906.8; DB 3; Length 1058;
Best Local Similarity 98.0%; Pred. No. 4e-244; Indels 12; Gaps 12;
Matches 1037; Conservative 5; Mismatches 4;
QY 98 GTCGCGCTGTGCCCCCGCGCATTTCTCTCCCGCTGCTT-CCTTCTCTCTGGGCTTC 156
Db 1 GCTGCGCTGTGCCCCCGCGCATTTCTCTCCCGCTGCTTCCCTTCTCTCTGGGCTTC 60
QY 157 GTGCTCTCTCGCTCCGATGGCGGAGCGGCTGCAACAA-GGCGCCCTTCCCTG 215
Db 61 CTGCTCTCTCGCTCCGATGGCGGAGCGGCTGCAACAAAGGCGGCTTCCCTG 120
QY 216 GATAGGCTCACTTTCTCAAGGTCAATCCCAAGCAAGTTCGTC-TGGTGAAGTTCAC 274
Db 121 GATAGGCTCACTTTCTCAAGGTCAATCCCAAGCAAGTTCGTC-TGGTGAAGTTCAC 180
QY 275 ACCAGTACCTTACGCTGAGAGCAGGATGAGTTCAGACGCTTT-CTGAAACCTCGGCT 333
Db 181 ACCAGTACCTTACGCTGAGAGCAGGATGAGTTCAGACGCTTTCTGTTGAAACCTCGGCT 240
QY 334 TCCAGCGATGATCTCTTGGTGCGAGAGTGGGATCTCAGATTAT-TGTGCAAGCTGAAC 392
Db 241 TCCAGCGATGATCTCTTGGTGCGAGAGTGGGATCTCAGATTAT-TGTGCAAGCTGAAC 300
QY 393 ATGGAGCTGAGTGAGAAATACAGCTGCAACAGAGAGCTACCCA-TCTTCTACCTCTTC 451
Db 301 ATGGAGCTGAGTGAGAAATACAGCTGCAACAGAGAGCTACCCAAGTCTTCTACCTCTTC 360
QY 452 CGGATGGGACTTTGAGAACCCAGTCCATACACTGGGGCAGTT-AGTTGGAGCCATC 510
Db 361 CGGATGGGACTTTGAGAACCCAGTCCATACACTGGGGCAGTT-AGTTGGAGCCATC 420
QY 511 CAGCGCTGGTGAAGGGCAAGGGGTCTACCTAGGTATGCCCTGG-TGCCCTGCTGTATAC 569
Db 421 CAGCGCTGGTGAAGGGCAAGGGGTCTACCTAGGTATGCCCTGG-TGCCCTGCTGTATAC 480
QY 570 GAGCGCTGGCGGGGAGTTCTACAGGCTCTGTGTGGAGG-CCGCGAGGCTCTTTC 628
Db 481 GAGCGCTGGCGGGGAGTTCTACAGGCTCTGTGTGGAGG-CCGCGAGGCTCTTTC 540
QY 629 AAGCAGGGGCAAGATTAACCTCTCAAGTGTGAAGAGACTCAAG-AGTGGGCGGAGCAA 687
Db 541 AAGCAGGGGCAAGATTAACCTCTCAAGTGTGAAGAGACTCAAG-AGTGGGCGGAGCAA 600
QY 688 TACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGCATCAGATG 746
Db 601 TACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTTCCAGCATCAGATG 660
QY 747 ACACGGATCGCCAGGCTGATTGAGAAGAAACAGATGAGTGAC-GGCAGAGGAGGAGCTC 805
```

```
Db 661 ACAGGATCGCCAGGCTGATTGAGAGAACAGATGAGTACGCGGAAGAGGAGGCTC 720
QY 806 CAGAAGAGCTTAAACATCTCCTGACTGCCCTTCAGAGAAAGGGGGCCGAGAGAGCTG 865
Db 721 CAGAAGAGCTTAAACATCTCCTGACTGCCCTTCAGAGAAAGGGGGCCGAGAGAGCTG 780
QY 866 TAAAGAGCTCTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGGGAGTTAACT 925
Db 781 TAAAGAGCTCTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGGGAGTTAACT 840
QY 926 GCTGCTGTGANTCCCTTGTGGAATATAAGGGGTMKGGGAAAGMGGTACTTAACCCAC 985
Db 841 GCTGCTGTGAGTCCCTTGTGGAATATAAGGGGTAGTGGGAAAGTGTACTTAACCCAC 900
QY 996 GATTCTGAGCCCTGAGTATGCTTGACATTCATGATCTAAACATGACCATCTTGGGATGCT 1045
Db 901 GATTCTGAGCCCTGAGTATGCTTGACATTCATGATCTAAACATGACCATCTTGGGATGCT 960
QY 1046 CTAGCTGCTGGGGATAGCTGGAGCACTTACTCAGGTGGCTGTGGAATGACACCTCAG 1105
Db 961 CTAGCTGCTGGGGATAGCTGGAGCACTTACTCAGGTGGCTGTGGAATGACACCTCAG 1020
QY 1106 AAGGAATGAGTGTCTATAGAGAGGAGAGAGGAGTGTACT 1143
Db 1021 AAGGAATGAGTGTCTATAGAGAGGAGAGAGGAGTGTACT 1058
RESULT 13
BX403041/c
LOCUS
DEFINITION
BX403041 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL010YA17 3-PRIME, mRNA sequence.
ACCESSION
BX403041
VERSION
BX403041.2 GI:46835608
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1098)
Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
On May 13, 2003 this sequence version replaced gi:30607274.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6716.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS1AL003ZC09NP1&c=6716.f.
Location/Qualifiers
1..1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL010YA17"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notice="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

| | | | | | |
|-----------------------|--------------|---|----------------|--------------|----------|
| Query Match | 75.1%; | Score 904.4; | DB 5; | Length 1098; | |
| Best Local Similarity | 95.1%; | Pred. No. 1.9e-243; | | | |
| Matches 1041; | Conservative | 25; | Mismatches 16; | Indels 13; | Gaps 13; |
| QY | 83 | GAACCGCGGATATGCTCCGCTGTGCCGCGCGCATTTCTCTCCCGCTGCTT-CC | 141 | | |
| DB | 1095 | GRAMCCGGDTTATTGSGTGCCSTTTKTCGCCSCCCCGCTWTTTCTCCCGCTGCTTCCC | 1036 | | |
| QY | 142 | CTTCTCTCGGCTTCCGCTCC-TCCTCGCTCCGATGGCGCGCAGCGGCTTGACACCAA | 200 | | |
| DB | 1035 | CTTCTCTCGGCTTCCGCTCTCTCTCCSTCCGATGCGCGCGGCTTGACACCAA | 976 | | |
| QY | 201 | -GGCGCCCTTCCCTCGGATACGGTCACTTTCTACAAGTCAATTCCTCCAAAGCAAGTTGCT | 259 | | |
| DB | 975 | GGCGCCCTTCCCTCGGATACGGTCACTTTCTACAAGTCAATTCCTCCAAAGCAAGTTGCT | 916 | | |
| QY | 260 | C-TGGTGAAGTTCGACACCGATCCCTACGGTTCAGAGCAGGATGAGTTCAGCGCTCT | 318 | | |
| DB | 915 | CTTGGTGAAGTTCACACCGATCCCTACGGTTCAGAGCAGGATGAGTTCAGCGCTCT | 856 | | |
| QY | 319 | T-CTGAAAACTCGGCTTCCAGCGATGATCTCTTGTGTGCAGAGGTGGGGATCTCAGATTA | 377 | | |
| DB | 855 | TGCTGAAAACTCGGCTTCCAGCGATGATCTCTTGTGTGCAGAGGTGGGGATCTCAGATTA | 796 | | |
| QY | 378 | T-GTGAACAGCTGAACATCGAGCTCAGTGAGAAATACAAAGTGGACAAAGAGACTACCC | 436 | | |
| DB | 795 | TGSTRACAAAGCTGAACATCGAGCTCAGTGAGAAATACAAAGTGGACAAAGAGACTACCC | 736 | | |
| QY | 437 | A-TCTTCTACCTTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGT | 495 | | |
| DB | 735 | AGTCTTCTACCTTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGT | 676 | | |
| QY | 496 | T-AGGTTGAGCCATCCAGCGCTGCTGAAGGGGCAAGGGGTCTACCTAGTATGCTCTGG | 554 | | |
| DB | 675 | TAAAGTTGAGCCATCCAGCGCTGCTGAAGGGGCAAGGGGTCTACCTAGTATGCTCTGG | 616 | | |
| QY | 555 | -TGCTTCCTGTATACGACGCCCTTGGCGGGAGTTATCAGGGGCTCTGCTGTGGAGG- | 612 | | |
| DB | 615 | TTSCCTGCCGTTATACGACGCCCTTGGCGGGAGTTATCAGGGGCTCTGCTGTGGAGG | 556 | | |
| QY | 613 | CCGCCAGGCCCTTTGAACACAGGGGCAAGATAACCTCTCAAGTGTGAAGGACTCAGAA | 672 | | |
| DB | 555 | CCGCCAGGCCCTTTGAACACAGGGGCAAGATAACCTCTCAAGTGTGAAGGACTCAGAA | 496 | | |
| QY | 673 | G-AGTGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT | 731 | | |
| DB | 495 | GAAGTGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT | 436 | | |
| QY | 732 | -CCAGCATCAGAGATGACACGGATCCGACGGTGTATTGAGAAGAACAGATGAGTGAC-G | 789 | | |
| DB | 435 | CCAGCATCAGAGATGACACGGATCCGACGGTGTATTGAGAAGAACAGATGAGTGACGG | 376 | | |
| QY | 790 | GCAGAAGGAGAGCTCCAGAAGACTTAAACATCTCTGACTGCTTCCAGAAGAGGGGGC | 849 | | |
| DB | 375 | GAAGAAGGAGAGCTCCAGAAGACTTAAACATCTCTGACTGCTTCCAGAAGAGGGGGC | 316 | | |
| QY | 850 | CGAGAAGGAGAGCTTAAAGAGGCTGTGTGATTTTCCAGGGTTTGGTGGGGGTAGGG | 909 | | |
| DB | 315 | CGAGAAGGAGAGCTTAAAGAGGCTGTGTGATTTTCCAGGGTTTGGTGGGGGTAGGG | 256 | | |
| QY | 910 | AGGGGANAGTTAACTCTGGCTTGANTCCCTTGTGGAAATATAAGGGGGMKSGGAA | 969 | | |
| DB | 255 | AGGGGAGAGTTAACTCTGGCTGTGTGATTTTCCAGGGTTTGGTGGGGGTAGTGGGAA | 196 | | |
| QY | 970 | AGWGGTACTAAACCCAGATTTGAGCCCTGAGTATGCTTGGACATGATGATTAACATGAC | 1029 | | |
| DB | 195 | AGTGGTACTAAACCCAGATTTGAGCCCTGAGTATGCTTGGACATGATGATTAACATGAC | 136 | | |
| QY | 1030 | CATGCTTGGGATGCTCTAGCTGGTCTGGGGATAGCTGGACACTTACTCAGGTGGCTGG | 1089 | | |
| DB | 135 | CATGCTTGGGATGCTCTAGCTGGTCTGGGGATAGCTGGACACTTACTCAGGTGGCTGG | 76 | | |
| QY | 1090 | TGAATGACACCTTCAGAAGGATAGTGCTATAGAGAGAGAGAGAGTGTACTGCCCCAG | 1149 | | |

| | | | |
|----|------|---|------|
| Qy | 324 | AAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGGATCTCAGATTAT-GTGA | 382 |
| Db | 241 | AAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGGATCTCAGATTATGTGTGA | 300 |
| Qy | 383 | CAAGCTGAACATGGAGCTGAGTGTAGAAATACAAGCTGTGCACAAAGAGACTACCCA-TCCTT | 441 |
| Db | 301 | CAAGCTGAACATGGAGCTGAGTGTAGAAATACAAGCTGTGCACAAAGAGACTACCCAGTCTT | 360 |
| Qy | 442 | CTACCTCTTCCGGGATGGGACTTTTGAGAACCCAGTCCCACATACACTGGGGCAGTTT-AGGT | 500 |
| Db | 361 | CTACCTCTTCCGGGATGGGACTTTTGAGAACCCAGTCCCACATACACTGGGGCAGTTTAAAGT | 420 |
| Qy | 501 | TGGAGCCATCCAGCGCTGGCTGTGAAGGGCAAGGGTCTACCTAGTATGCCTGG-TGGCT | 559 |
| Db | 421 | TGGAGCCATCCAGCGCTGGCTGTGAAGGGCAAGGGTCTACCTAGTATGCCTGGTGGCT | 480 |
| Qy | 560 | GCCGTGTATACGACGCCCTTGGCCGGGGAGTTCATCAGGGCCCTCTGGTGTGGAGG-CCGCCA | 618 |
| Db | 481 | GCCGTGTATACGACGCCCTTGGCCGGGGAGTTCATCAGGGCCCTCTGGTGTGGAGGCCGCCA | 540 |
| Qy | 619 | GGCCCTCTTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAG-AGTG | 677 |
| Db | 541 | GGCCCTCTTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAGAAGTG | 600 |
| Qy | 678 | GGCCGAGCAATACCTGAAGATCATTGGGGAAGATCTTAGACCAAGGGGAGACATT-CCAGC | 736 |
| Db | 601 | GGCCGAGCAATACCTGAAGATCATTGGGGAAGATCTTAGACCAAGGGGAGGACTTCCCAGC | 660 |
| Qy | 737 | ATCAGAGATGACCGGATCGCCAGGCTGATTGTAGAAGAAACAAGATGAGTGTAC-GGCAGAA | 795 |
| Db | 661 | ATCAGAGATGACCGGATCGCCAGGCTGATTGTAGAAGAAACAAGATGAGTGTACGGGAAGNA | 720 |
| Qy | 796 | GGAGGAGCTCCAGAAGAGCTTTAAACATCCTGACTGCCTTCAGAGAAGGGGGCCCGAGAA | 855 |
| Db | 721 | GGAGGAGCTCCAGAAGAGCTTTAAACATCCTGACTGCCTTCAGAGAAGGGGGCCCGAGAA | 780 |
| Qy | 856 | AGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTTCAGGGTTTGGTGGGGGTAGGAGGGGA | 915 |
| Db | 781 | AGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTTCAGGGTTTGGTGGGGGTAGGAGGGGA | 840 |
| Qy | 916 | NAGTTTAACTGCTCGCTGTGANTCCCTTGTGGAATATAAGGGGGGWMKGGGCAAAAGWGT | 975 |
| Db | 841 | GAGTTTAACTGCTCGCTGTGAGTCCCTTGTGGAATATAAGGGGGGTAGTGGGAAAGTGGT | 900 |
| Qy | 976 | ACTAACCACGATTTCTGAGCCCTCAGTATGCCTGGAACATTGATGCTAACATGACCATGCT | 1035 |
| Db | 901 | ACTAAMCCACGATTTCTGAGCCCTCAGTATGCTGTGGAACATTGATGCTAACATGACATGCT | 960 |
| Qy | 1036 | TGGGATGTCTTACGTGGTCTTGGGGAATAGCTGGAGCACTTACTCAGGTGGCTGTGTGAA | 1093 |
| Db | 961 | TGGGATGTCTTACGTGGTCTTGGGGAAT-STGGAGCACTTACTCAGGTGGCTGTGTGAA | 1017 |

RESULT 15

[illegible]

| | | | |
|----|------|--|------|
| Qy | 781 | TGAGTGAC--GGCAGAGGAGGAGCTCCAGAAAGAGCTTAAACATCTCTGATGCTCCCTCCAGA | 839 |
| Db | 409 | TGAGTGACGGGAAAGAGGAGGAGCTCCAGAAAGAGCTTAAACATCTCTAACTGCTCCCTCCAGA | 350 |
| Qy | 840 | AGAAGGGGGCCGAGAAGAGAGAGCTGTAAAAAAGCTGCTGTGATTTTCCAGAGTTTGGT | 899 |
| Db | 349 | AGAAGGGGGCCGAGAAAGAGAGAGCTGTAAANAAGCTGCTGTGATTTTCCAGGGTTGGT | 290 |
| Qy | 900 | GGGGGTAGGGAGGGGAGNAGTTAACTGCTGGCTGTGANTCCCTGTGGAATATAAGGGGG | 959 |
| Db | 289 | GGGGGTAGGGAGGGGAGAGTAACCTGCTGGCTGTGAGTCCCTGTGTGGAATATAAGGGGG | 230 |
| Qy | 960 | YMSKGGGAAAGHWGFTACTAACCCACGATTCTGAGCCCTGAGTGATGCTGGGACATTTGATG | 1019 |
| Db | 229 | TAGTGGGAAAGTGGTACTAACCCACGATTCTGAGCCCTGAGTGATGCTGGGACATTTGATG | 170 |
| Qy | 1020 | CTAACATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGGATAGCTGGAGGACATTTACTC | 1079 |
| Db | 169 | CTAACATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGGATAGCTGGAGGACATTTACTC | 110 |
| Qy | 1080 | AGGTGGCTGCTGAAATGACACCTCAGAAGGAAATGAGTGTCTATAGAGAGGAGAGAGGATG | 1139 |
| Db | 109 | AGGTGGCTGCTGAAATGACACCTCAGAAGGAAATGAGTGTCTATAGAGAGGAGAGAGGATG | 50 |
| Qy | 1140 | TACTGCCCCAGGT--CTTTGACAGATGTAATCTCAATTCAAATTAAGTTTT | 1186 |
| Db | 49 | TACTGCTNCAGKTCCTTTTATCAGGTGTAATCTCTCNNTCDMGTTAAHTTT | 1 |

Search completed: July 3, 2005, 16:39:54
Job time : 3817.04 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 07:04:57 ; Search time 10073.7 Seconds
(without alignments)
12038.564 Million cell updates/sec

Title: US-08-731-499-10
Perfect score: 3186
Sequence: 1 ATGCAATCGAAGTACACAGG.....GGACACTACAGTTGTGTAA 3186

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|------------|--------------|
| 1 | 1028.6 | 32.3 | 2331 | 3 AK030249 | Mus musculus |
| 2 | 762.2 | 23.9 | 841 | 6 CD644465 | AGENCOURT |
| 3 | 647 | 20.3 | 651 | 5 BX119280 | BMX119280 |
| 4 | 615.6 | 19.3 | 618 | 5 BX955300 | DKFZp781N |
| 5 | 590 | 18.5 | 595 | 4 BF944766 | PM3-NN117 |
| 6 | 581.4 | 18.2 | 583 | 5 BP336988 | BP336988 |
| 7 | 471.8 | 14.8 | 569 | 2 AW753609 | RC1-CT026 |
| 8 | 470.8 | 14.8 | 895 | 9 CR094054 | CR094054 |
| 9 | 466 | 14.6 | 479 | 1 A1380386 | Forward |
| 10 | 453.2 | 14.2 | 805 | 6 CB522302 | UI-N-GH0- |
| 11 | 424.4 | 13.3 | 692 | 2 BB617224 | BB617224 |
| 12 | 418.8 | 13.1 | 768 | 2 BF237807 | BF237807 |
| 13 | 395.2 | 12.4 | 473 | 7 W05407 | W05407 |
| 14 | 394 | 12.4 | 411 | 4 BG007233 | IL5-GN023 |
| 15 | 384 | 12.1 | 640 | 4 BM834501 | K-EST0109 |
| 16 | 377.4 | 11.8 | 594 | 7 CV409050 | RC1-HT026 |
| 17 | 376.2 | 11.8 | 636 | 6 CB457791 | 715904 MA |
| 18 | 369.2 | 11.6 | 519 | 1 AI025334 | ov74905.x |
| 19 | 369 | 11.6 | 399 | 2 BB819616 | RC2-BN033 |
| 20 | 368.6 | 11.6 | 400 | 5 BQ367562 | IL5-GN023 |
| 21 | 365.4 | 11.5 | 368 | 4 BM146822 | TCAAP1874 |
| 22 | 350.8 | 11.0 | 412 | 1 AA069107 | zm10b10.x |
| 23 | 347.4 | 10.9 | 703 | 4 B1185557 | UNL-P-PN- |
| 24 | 346.2 | 10.9 | 353 | 4 BG013874 | IL5-GN023 |

| | | | | | |
|------|-------|------|-----|------------|------------|
| C 25 | 344.6 | 10.8 | 662 | 6 CB460210 | CB460210 |
| C 26 | 343.2 | 10.8 | 440 | 7 N70546 | zab3c12.81 |
| C 27 | 318 | 10.0 | 674 | 6 BY748714 | BY748714 |
| C 28 | 309.4 | 9.7 | 667 | 8 A2939266 | A2939266 |
| C 29 | 301.8 | 9.5 | 305 | 4 BG982616 | IL5-CN006 |
| C 30 | 300 | 9.4 | 352 | 4 BF963063 | PM3-NN117 |
| C 31 | 299 | 9.4 | 328 | 2 BB818091 | CM2-BN027 |
| C 32 | 294.4 | 9.2 | 571 | 7 CR753532 | DKFZp469L |
| C 33 | 284.2 | 8.9 | 576 | 2 BF782572 | 602106595 |
| C 34 | 282 | 8.9 | 327 | 2 BB817937 | CM2-BN027 |
| C 35 | 278.2 | 8.7 | 848 | 1 AJ743843 | AJ743843 |
| C 36 | 264.4 | 8.3 | 313 | 2 BF803316 | RC1-CI011 |
| C 37 | 262.6 | 8.2 | 284 | 2 BB819568 | RC2-BN033 |
| C 38 | 260.4 | 8.2 | 268 | 4 BG012719 | IL5-GN023 |
| C 39 | 258 | 8.1 | 308 | 2 BB818055 | CM2-BN027 |
| C 40 | 256.4 | 8.0 | 341 | 2 BF329627 | CM2-BN027 |
| C 41 | 254 | 8.0 | 332 | 4 BG978004 | PM1-CI015 |
| C 42 | 251.4 | 7.9 | 254 | 1 AU185193 | AU185193 |
| C 43 | 248.6 | 7.8 | 339 | 4 BG012812 | IL5-GN023 |
| C 44 | 247.2 | 7.8 | 749 | 1 AJ743855 | AJ743855 |
| C 45 | 246.6 | 7.7 | 262 | 2 AW380095 | QV3-HT026 |

ALIGNMENTS

| | | | | | |
|------------|---|---------|------|--------|-----------------|
| RESULT 1 | AK030249 | 2331 bp | mRNA | linear | HTC 03-APR-2004 |
| LOCUS | AK030249 | | | | |
| DEFINITION | Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4933431C08 product:weakly similar to ZINC FINGER PROTEIN 217 [Homo sapiens], full insert sequence. | | | | |
| ACCESSION | AK030249 | | | | |
| VERSION | AK030249.1 | | | | GI:26081731 |
| KEYWORDS | HTC; CAP trapper. | | | | |
| SOURCE | Mus musculus (house mouse) | | | | |
| ORGANISM | Mus musculus | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | |
| AUTHORS | Carninci, P. and Hayashizaki, Y. | | | | |
| TITLE | High-efficiency full-length cDNA cloning | | | | |
| JOURNAL | Meth. Enzymol. 303, 19-44 (1999) | | | | |
| MEDLINE | 99279253 | | | | |
| PUBMED | 10349636 | | | | |
| REFERENCE | 1 | | | | |
| AUTHORS | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | | | | |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | | | | |
| JOURNAL | Genome Res. 10 (10), 1617-1630 (2000) | | | | |
| MEDLINE | 20499374 | | | | |
| PUBMED | 11042159 | | | | |
| REFERENCE | 3 | | | | |
| AUTHORS | Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. | | | | |
| TITLE | RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer | | | | |
| JOURNAL | Genome Res. 10 (11), 1757-1771 (2000) | | | | |
| MEDLINE | 20530913 | | | | |
| PUBMED | 11076861 | | | | |
| REFERENCE | 4 | | | | |
| AUTHORS | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. | | | | |
| TITLE | Functional annotation of a full-length mouse cDNA collection | | | | |
| JOURNAL | Nature 409, 685-690 (2001) | | | | |
| PUBMED | 11855557 | | | | |
| REFERENCE | 5 | | | | |
| AUTHORS | The FANTOM Consortium and the RIKEN Genome Exploration Research | | | | |

Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2331)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gsr.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://phantom.gsc.riken.jp/
Location/Qualifiers
1..2331
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:4933431C08"
/db_xref="taxon:10090"
/clone="4933431C08"
/sex="male"
/tissue type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
228..2104
/note="putative
weakly similar to ZINC FINGER PROTEIN 217 [Homo sapiens]
(SWISSPROT|O75362, evidence: FASTV, 69.5%ID, 67.8%length,
match=2090)"]
misc_feature
ORIGIN
Query Match 32.3%; Score 1028.6; DB 3; Length 2331;
Best Local Similarity 72.4%; Pred. NO. 1.1e-286;
Matches 1480; Conservative 0; Mismatches 504; Indels 59; Gaps 9;
QY 1004 AAAAATAAGGCGAGTGTGCGAGCCCTCTCGCAAGAGAAAGAGAGTGCACACTCCC 1063
DB 4 AAAAAGTAAAGCAGTGTGTCAGGTCTCTCCCAAGCAAGAGAGAGCTAGATGCTA 63
QY 1064 ACGGCGAAGCGCCCTCCGTCGACGGGATCCCAAGTTACCCAGTAGCAAGAGAGCCCA 1123
DB 64 ATATGTAAGTGCCTCTGGGNTAGTAGCCCNAGTTGTCNAGTAGCAAGAGAGCCCA 123
QY 1124 CTCAGTCTCCGAGTGGCGAAAGTTTCAGAACCTACACAGTGTCTTGCACTCCA 1183
DB 124 CGCACTGTCTAGTGCAGCAAGCTTTCAGACATACACAGCTCGTCTGCACTCGA 183
QY 1184 GGGTCCACAGAGACCGAGGGCGGCGGAGTCCGCCACCATGCTCTGTGACGGGA 1243
DB 184 GGGTGCAAGGAGACAGGAGGACTGATGCCCTGTCAACCCACCATGCTGTGGATGCA 243
QY 1244 GGCAGCCGGGACGTGTTCTCCTGACCTCGCGCCCTCTGGATGAAAAATGGAGCCGTGG 1303
244 GGCAGCCCTGGGACCTGCTGCCAGACCTCAGCACCACTCTGGAAGACAGTGGGGCCGGG 303
1304 ATCGAGGGGAAGTGGTCTTGAAGACGGATCTGAGGATGGCTTCCCGAAGGATCCATC 1363
304 ACCCA---GAAGGGGGCTCTGAAGACGGGTCTGAGGATGGACTCCCTGAGCGGGTCCATT 360
1364 TGGATAAAATGATGATGAGGAGGAAAAATAA--ACATCTTACATCTTCAAGAGAGTGTAG 1421
361 TGGNTAAATGATGATGAGGAGGAAAAAGGAAAGCCCTCCCTCCCTCGAGAGAGTGTAG 420
1422 TTATTTGGAAGTTTTCGGTTCAAATTAATTAACCTCAATTAATTCATCTCAGAAACGCTATC 1481
421 TTACTGTGGCAAGTTTTCGGTTCAAACTATTACCTCAATTAATTCATCTCAGAAACGCTATC 480
1482 AGGTGAAAAACCATACAAATGTGAAATTTTGTGAATGATGTCGACGCCAGAGAGATCTCT 1541
481 AGGTGAAAAACCATACAAATGTGAAATTTCTGTGAGTATGCGCGACGCCAGAGAGATCTCT 540
1542 GAGGTATCACTTGGAGAGACATCAAGAGGAAAAACAAC---CGATGTTGCTGCTGAAGT 1598
541 GAGGTACCACTTGGAGAGACATCAAGAGACAGCAGCCGGTGTGATGCTGCCGTGAGTC 600
1599 CAAGAACGATGTTAAAAATCAGGACACTGAAGATGCACTATTAA---CCGCTGACAGTGC 1655
601 CAAAAGTGAAGCGCGGAGCCAGGAGCCGAGGATGCGCTACTAAACGGCTGCTGACAGTGC 660
1656 GCAACCAAAAATTTGAAAAGATTTTGTGATGGGCCAAGAGTGTTCAGAGCATGTCACC 1715
661 GCAGACCAAAAATTTAAAGAGATTTCTGTGATGGTGCACAAAGATGTTAAGGAAAGCCACC 720
1716 TGCAAGCAGCTTAAGGAGATGCTCTGTTTTTTCAGAAATGTTCTGGGCGCAGCTGTCCT 1775
721 TGCCAGCAGCTTAAGGAGATGCTCTGTTTTCCAGAGTGTTC-----TCTC 768
1776 CTCACAGCACACAAAGATCTCAGGATTTCCATAAAATGCACTGATGACAGTGTGTA 1835
769 ACCAGCACACAGCAACGATCTCAGGATTTCCATAAAACATGACAGTGTAT---AGTGTGA 825
1836 TAAAGTGAATAAAACCCCTACCCCTGCTTACCTGGACCTGTTTAAAGAGAGATCAGCAGT 1895
826 GAAAGCGAGGAGAGAGCCCTGCCCCCTACTTATCTGGACATGACAGAGAAAGA---AAGCAG 882
1896 TGAACCTCAGGCAATAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCCGGA 1955
883 GGACCTTCAGGCCAGAGCCCTGCTGACAGACTAGAGGGGTTGGTCTCTTAGCACGGA 942
1956 TGGCAGTACCAACCCTAACCTTGAAGTTAGCCCCCAAGAGAGCAAGCGGAGACCGCAGC 2015
943 AGCTGG-----CCATAGGGAGAGATGATGATCAGGATGC 975
2016 TGACTGCGATACAGGCCCAAGTGTGGATTTGTCAGAAAAACCTTTTAAATTTATCCGTGGG 2075
976 TGACTACAGACATAAGCCCCGGTGTGACTGCGCAGACAGGCCCTTTGAATCTATCCCTTGG 1035
2076 GGCTCTTCACAATTTGCCCGGCAATTTCTTTGAGTAAAGTTTGTATTCACAGTATCACCTG 2135
1036 GCGCTCCACGCTGTCTTGCAATCTCTTTGAGCAAGTGTCTGATCCCCASCATTCGCTG 1095
2136 TCCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTTTAATGATGACACAGAGACTGGA 2195
1096 CCCCTTTTGTACTTTCAAGACCTTTTATCCGAAAGTCTTATGATGACACAGAGACTTGA 1155
2196 GCATAATATCAATCTCAGCTTCAATAAACTTTCGAAACCAAGTCTTGTAGAGTCG 2255
1156 GCACAGGTACAAACCTGACCCGCAAGAACCGGACAGCAAGTCTGTGCTGAGAGAACAG 1215
2256 ACGTACCGGATGCCCGCAGCGTTGCTGGGAAAAAGATGTGCTCCCTCTCTAGTTTCTG 2315
1216 GCGTACCGGGTGCCTCGGCTTTGCTGGGAAAGAGTGTCCCTCTCTGCTGGCCTGCA 1275
2316 TAAACCCAGCCCAAGTGTGCTTTCCGCGGCGAGTCCAAATTCCTCTGCCATCTGCGAAGGG 2375


```
Db 241 CACCAGGCAACAGCAATCTGAGATGTTTCTTAAACACAGTGTTCCTCCCTGACCGGATAA 300
Qy 2559 GACAAAAGAGCCCGAGACAAAATTAACACCTTCCAGTAGCTCTTCTCAGCCACCCCT 2618
Db 301 GACAAAAGAGCCCGAGACAAAATTAACACCTTCCAGTAGCTCTTCTCAGCCACCCCT 360
Qy 2619 CGGAGCAGTAGTAACATCAATAGTGTTCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGC 2678
Db 361 CGGAGCAGTAGTAACATCAATAGTGTTCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGC 420
Qy 2679 ACCTCCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATATCTGAGCAGAAATTTGG 2738
Db 421 ACCTCCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATATCTGAGCAGAAATTTGG 480
Qy 2739 TGAGCCCTTCCAAAAGAGACTGAAGTCCAGCGTGGTTCGCTTCAAGTTGACCAAGCCCGG 2798
Db 481 TGAGCCCTTCCAAAAGAGACTGAAGTCCAGCGTGGTTCGCTTCAAGTTGACCAAGCCCGG 540
Qy 2799 GGCCAAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGGCATCACATC 2858
Db 541 GGCCAAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGGCATCACATC 600
Qy 2859 ACTGTTACCGCAGGACTGTGTATTCCTGTCGAGCGCTGCTCCCAAACCAAGGTTCT 2918
Db 601 ACTGTTACCGCAGGACTGTGTATTCCTGTCGAGCGCTGCTCCCAAACCAAGGTTCT 660
Qy 2919 GAGTCCAGCAGGCTCGATTTCCAAATGTCGTGACTGTTCCAGAGCCCTATGTGGCTC 2978
Db 661 GAGTCCAGCAGGCTCGATTTCC - AATGTCTGACTGTTCCAGAAAGCCTATGTGGCTC 719
Qy 2979 CGGGCCACTTTACACTTGTGTGCTGCTGTGTAGTCCAGCATCCAGCTCGACGTTAGAGG 3038
Db 720 CGGGCCACTTTACANCTGTGTGCTGCTGTGTAGTCCAGCATCCAGCTCGACGTTAGAGG 779
```

```
RESULT 3
BX119280
LOCUS
DEFINITION BX119280 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
IMAGp998F23669 ; IMAGE:299158, mRNA sequence.
ACCESSION BX119280
VERSION
KEYWORDS EST.
SOURCE BX119280.1 GI:27882834
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ebert L., Heil O., Hennig S., Neubert P., Partsch E., Peters M.,
Radelof U., Schneider D. and Korn B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rols
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGp998F23669,
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rols
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.
LOCATION/Qualifiers
1..651
/organism="Homo sapiens"
/mol_type="mRNA"
```

FEATURES
source

RESULT 4
BX955300
LOCUS
DEFINITION DXFzp781N0270_r1 781 (synonym: hlccc4) Homo sapiens cDNA clone
DXFzp781N0270 5', mRNA sequence.

```
/db xref="taxon:9606"
/clone="IMAGp998F23669 ; IMAGE:299158"
/dev stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_lung_NbHL19W"
/notes="Organ: lung; Vector: p7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGGCCCAATTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Patima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NbHL19W."
```

ORIGIN

| Query Match | 20.3% | Score 647; | DB 5; | Length 651; |
|-----------------------|---|---------------------|---------------|-------------------|
| Best Local Similarity | 99.5% | Pred. No. 6.3e-176; | | |
| Matches | 647; | Conservative 0; | Mismatches 3; | Indels 0; Gaps 0; |
| Qy 1107 | TAGCAAGAGAGCCCACTCCTCCGAGTGGCGCAAGCTTTTCAGAACCTTACCACCA | 1166 | | |
| Db 1 | TAGCAAGAGAGAGCCCACTCCTCCGAGTGGCGCAAGCTTTTCAGAACCTTACCACCA | 60 | | |
| Qy 1167 | GCTGCTCTGCACCTCCAGGGTCCACAAGAGAGCCGCGGCGCGGAGTCCGCCAC | 1226 | | |
| Db 61 | GCTGCTCTGCACCTCCAGGGTCCACAAGAGAGCCGCGGCGCGGAGTCCGCCAC | 120 | | |
| Qy 1227 | CATGCTGTGGAGCGGAGCGAGCGGGAGCGTTCCTCCTGACCTCGCGCCCTCTGGA | 1286 | | |
| Db 121 | CATGCTGTGGAGCGGAGCGAGCGGGAGCGTTCCTCCTGACCTCGCGCCCTCTGGA | 180 | | |
| Qy 1287 | TGAAAATGGAGCCGTGGATCGAGGGAGAGTGGTTCCTGAAGACGGATCTGAGGATGGCT | 1346 | | |
| Db 181 | TGAAAATGGAGCCGTGGATCGAGGGAGAGTGGTTCCTGAAGACGGATCTGAGGATGGCT | 240 | | |
| Qy 1347 | TCCCGAAGGAATCCATCTGGATAAAATGATGAGGAGGAGAAATAAACAATCTTACATC | 1406 | | |
| Db 241 | TCCCGAAGGAATCCATCTGGATAAAATGATGAGGAGGAGAAATAAACAATCTTACATC | 300 | | |
| Qy 1407 | TTCAAGAGAGTGTAGTTATTGTGAAAGTTTTTCCGTTCAAATTTATTAATTTCA | 1466 | | |
| Db 301 | TTCAAGAGAGTGTAGTTATTGTGAAAGTTTTTCCGTTCAAATTTATTAATTTCA | 360 | | |
| Qy 1467 | TCTCAGAACCGATACAGGTGAAAACCATACAAATGTGAATTTTGTGAATATGCTGAGC | 1526 | | |
| Db 361 | TCTCAGAACCGATACAGGTGAAAACCATACAAATGTGAATTTTGTGAATATGCTGAGC | 420 | | |
| Qy 1527 | CCAGAAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAGAAAACCAACCGATGT | 1586 | | |
| Db 421 | CCAGAAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAGAAAACCAACCGATGT | 480 | | |
| Qy 1587 | TGCTGCTGAAGTCAAGAACCGATGTTAAATAATCAGGACATCAAGATGCATTTAAACCGC | 1646 | | |
| Db 481 | TGCTGCTGAAGTCAAGAACCGATGTTAAATAATCAGGACATCAAGATGCATTTAAACCGC | 540 | | |
| Qy 1647 | TGACAGTGGCGAAACCAAAAATTTGAAAAGATTTTTTGATGCTGCCAAGATGTTACAGG | 1706 | | |
| Db 541 | TGACAGTGGCGAAACCAAAAATTTGAAAAGATTTTTTGATGCTGCCAAGATGTTACAGG | 600 | | |
| Qy 1707 | CAGTCCACCTCGAAAGCAGCTTAAGGAGATGCCCTTCTGTTTTTCAGAAATG 1756 | | | |
| Db 601 | CAGTCCACCTCGAAAGCAGNTTAAGGAGATGCCCTTCTGTTTTTCANAATG 650 | | | |

BX955300 618 bp mRNA linear EST 01-MAR-2004
DXFzp781N0270_r1 781 (synonym: hlccc4) Homo sapiens cDNA clone
DXFzp781N0270 5', mRNA sequence.

[illegible]

```
Qy 357 GGAAGAAGATTGCAAGGAAAAATGAATTTAGCTGTGAGGTATGTGGCGACAGATTTAGAGT 416
Dy 61 GGAAGAAGATTGCAAGGAAAAATGAATTTAGCTGTGAGGTATGTGGCGACAGATTTAGAGT 120
Qy 417 CGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTTCTTTCATTAGGGGTGTAA 476
Dy 121 CGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTTCTTTCATTAGGGGTGTAA 180
Qy 477 CATGTGCGGAAGAAGMTTSSRSSAGCCTTGTTTCTTAAATAATCACATGCGGACACATAA 536
Dy 181 CATGTGCGGAAGAAGATTCAAGGAGCCTTGTTTCTTAAATAATCACATGCGGACACATAA 240
Qy 537 TGGCAAAATCGGGGGCGAGAACAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGAT 596
Dy 241 TGGCAAAATCGGGGGCGAGAACAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGAT 300
Qy 597 CAACGAGTGTCTGAGTGCACGCGGCGGAGAGATCTCTCTCTTACAAAATCTGAT 656
Dy 301 CAACGAGTGTCTGAGTGCACGCGGCGGAGAGATCTCTCTCTTACAAAATCTGAT 360
Qy 657 GGTGTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACAC 716
Dy 361 GGTGTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACAC 420
Qy 717 CAAAAAAACTGCTTTTCGTACGAGCGCGGCGGAGAGATCTCTCAAGGAGGAATGCC 776
Dy 421 CAAAAAAACTGCTTTTCGTACGAGCGCGGCGGAGAGATCTCTCAAGGAGGAATGCC 480
Qy 777 GTCTCTGAGGAGGAGCTTCTGCGAGTCTTCACTTGAGACAAAATCTCACCTGAAAC 836
Dy 481 GTCTCTGAGGAGGAGCTTCTGCGAGTCTTCACTTGAGACAAAATCTCACCTGAAAC 540
Qy 837 GGGGAAGAAGCCTGTGCAGATCATCCCTCAGCTCGATCCGTCACCACTTCCA 890
Dy 541 GGGGAAGAAGCCTGTGCAGATCATCCCTCAGCTCGATCCGTCACCACTTCCA 594
```

```
RESULT 6
LOCUS BP336988 Sugano cDNA library, coronary artery smooth muscle cell EST 17-SEP-2004
DEFINITION Homo sapiens cDNA clone SMR07185, mRNA sequence.
ACCESSION BP336988
VERSION BP336988.1 GI:52266581
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 583)
```

```
REFERENCE Suzuki Y., Yamaehita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
```

```
FEATURES
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SMR07185"
/tissue_type="coronary artery"
/cell_type="smooth muscle cell"
/clone_lib="Sugano cDNA library, coronary artery smooth
muscle cell"
```

ORIGIN

```
Query Match 18.2%; Score 581.4; DB 5; Length 583;
Best Local Similarity 99.8%; Pred. No. 7.1e-157;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1626 TGAAGATGCACTATTAACCGCTGACAGTGCAGCAACCAAAATTTGAAAAGATTTTGA 1685
Dy 1 TGAAGATGCACTATTAACCGCTGACAGTGCAGCAACCAAAATTTGAAAAGATTTTGA 60
Qy 1686 TGGTGCCAAAGATCTTACAGGAGCTCCACCTGCAAGCAGCTTAAGGAGATGCTCTCTGT 1745
Dy 61 TGGTGCCAAAGATCTTACAGGAGCTCCACCTGCAAGCAGCTTAAGGAGATGCTCTCTGT 120
Qy 1746 TTTTCAGAAATGTTCTGGGAGCGCTGCTCTCAACGACACAAAGATATCTCAGGATTT 1805
Dy 121 TTTTCAGAAATGTTCTGGGAGCGCTGCTCTCAACGACACAAAGATATCTCAGGATTT 180
Qy 1806 CCATAAAATGCAGCTGATGACAGTGTGATAAGTGAATAAAACCTCACCCCTGCTTA 1865
Dy 181 CCATAAAATGCAGCTGATGACAGTGTGATAAGTGAATAAAACCTCACCCCTGCTTA 240
Qy 1866 CCTGGACCTGTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATACCTCATCTGTAG 1925
Dy 241 CCTGSACTGTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATACCTCATCTGTAG 300
Qy 1926 AACCAAGCGGATTTACTCTCTCCGATGGCAGTACCAACCATACCTTGAAGTTAG 1985
Dy 301 AACCAAGCGGATTTACTCTCTCCGATGGCAGTACCAACCATACCTTGAAGTTAG 360
Qy 1986 CCCCAGAGAGAACAAACGAGACCGCAGTGTGAGATACAGGCAAGGAGTGGATTG 2045
Dy 361 CCCCAGAGAGAACAAACGAGACCGCAGTGTGAGATACAGGCAAGGAGTGGATTG 420
Qy 2046 TCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCTTCACAATTCGCCGCAATTTCTTT 2105
Dy 421 TCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCTTCACAATTCGCCGCAATTTCTTT 480
Qy 2106 GAGTAAAGTTTGATTCAGATATCACCTGTCCATTTTGTATCTTCAAGACATTTTATCC 2165
Dy 481 GAGTAAAGTTTGATTCAGATATCACCTGTCCATTTTGTATCTTCAAGACATTTTATCC 540
Qy 2166 AGAGTTTAAATGATGCACAGAGACTGGAGCATAAATACAAAT 2208
Dy 541 AGAAGTTTAAATGATGCACAGAGACTGGAGCATAAATACAAAT 583
```

```
RESULT 7
LOCUS AW753609/c
DEFINITION RCI-CT0268-060100-013-h08 CT0268 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW753609
VERSION AW753609.1 GI:7668541
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 569)
```

```
REFERENCE Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
```

```
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
```

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC1&t2=RC1-CT0268-060100-013-h08&t3=2000-01-06&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 49

High quality sequence stop: 67.

FEATURES

source

Location/Qualifiers

1..569

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="CT0268"

/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 14.8%; Score 471.8; DB 2; Length 569;

Best Local Similarity 92.1%; Pred. No. 5.2e-125;

Matches 524; Conservative 5; Mismatches 37; Indels 3; Gaps 3;

QY 113 TGTCATGAAGGGACCGCTGTGTTCCATTCGAGCTACACAGAAAGAAATGTCATCC 172

DB 569 TGTCATGAAGGGACCGCTGTGTTCCATTCGAGCTACACAGAAAGAAATGTCATCC 510

QY 173 AAATCGAGGGTATATGCGCTTGGATTGTCATGTTCTGCA-GCCAGACCTTTCACATTCAC 231

DB 509 CACTCGAGGGTATATGCGCTTGGATTGTCATGTTCTGCA-GCCAGACCTTTCACATTCAC 450

QY 232 GAAGACCTTAATAACATGTTTAATGCAACACCGCGCTTACCC-TCTGTGAACACGAGT 290

DB 449 GAAGACCTTAATAACATGTTTAATGCAACACCGCGCTTACCC-TCTGTGAACACGAGT 390

QY 291 TCTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATTAAGTCAAGTCCGAGACCA 350

DB 389 TCTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATTAAGTCAAGTCCGAGACCA 330

QY 351 TCCCAAGGAAAGAAATGCAAGAAATGAATTTAGTGTGAGTATGTGGCAGACATT 410

DB 329 TCCCAAGGAAAGAAATGCAAGAAATGAATTTAGTGTGAGTATGTGGCAGACATT 270

QY 411 TAGATCGCTTTTGTATGTTAGATCCATGAGAACACACAAAGATTCTTTTCACTTACGG 470

DB 269 TAGATCGCTTTTGTATGTTAGATCCATGAGAACACACAAAGATTCTTTTCACTTACGG 210

QY 471 GTGTAACATGCGGAAGAGMTTSSAGCTTGGTTTCTTAATAATCAGATCGGAC 530

DB 209 GTGTGACATGTGCGGTAGAAGATTGAAGAGCTTGGTTTCTTAATAATCAGATCGGAC 150

QY 531 ACATAATGCAAAATCGGGGCGAGAGCAAACTGCGAGCAAGCTTGGAGAGTAGTCCAGC 590

DB 149 ACATAATGCAAAATCGGGGCGAGAGCAAACTGCGAGCAAGCTTGGAGAGTAGTCCAGC 90

QY 591 AACGATCAACGAGGTGCTCGAGGTGACCGCGCGAGAGCATCTCTCTCC-TTACAAA 649

DB 89 CACGATCAACGAGGTGCTCGAGGTGACCGCGCGAGAGCATCTCTCTCC-TTACAAA 30

QY 650 TCTGATGTTTGTGCGTCTCTTATTTCCA 678

DB 29 TCTGATGTTTGTGCGTCTCTTATTTCCA 1

RESULT 8

CR094054

LOCUS

DEFINITION

Forward strand read from insert in 5'HPRT insertion targeting and

chromosome engineering clone MHPN19116, genomic survey sequence.

ACCESSION

CR094054

VERSION

CR094054.1

KEYWORDS

GSS; genome survey sequence; MICR.

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 895)

AUTHORS

Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,

Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,

Rogers,J., and Bradley,A.

TITLE

Direct Submission

JOURNAL

Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. http://www.sanger.ac.uk/MICR

FEATURES

Location/Qualifiers

1..895

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="MHPN19116"

/clone_lib="MHPN"

ORIGIN

Query Match 14.8%; Score 470.8; DB 9; Length 895;

Best Local Similarity 72.0%; Pred. No. 1.2e-124;

Matches 647; Conservative 6; Mismatches 239; Indels 7; Gaps 3;

QY 13 GTGACAGAAATGCGCAACTCAATCCCTCTTAATGTACATGATGGCCAGAAATGATT 72

DB 1 GTGGTGGCAGCATGCCAGCCAGTCCCTCTCTCGTGT-CATGACGGGCGGAAATGCTTC 59

QY 73 GGCAGCTCTTGGCAGTCCGATGAGATGAGGATGCTTGTCTAATGAAGGGACCGCT 132

DB 60 AGCAGCTCTTAGGCTCCAGATGAGGATGATGTGTGCCCATTAAGGGCGGTG 119

QY 133 GTTGTTCATTCGAGCTTACACAGAAAGAAATGTCTCAATCGAGGGGTATATGCC 192

DB 120 GCAGTCCCTTCCGAGTGTCTCAGGAGAGCAGTGGCGGAGAGGGCCACATGCC 179

QY 193 TTGATTCATGTTTTCGAGCAGACCTTACACATTCAGAAAGCCTTAATAAATGTC 252

DB 180 CTGATTCATGTTTTCGAGCAGACCTTCTCTCAGGGCGAGGATCTCAGTCAGCACG 239

QY 253 TTAATGCAACACCGGCTTACCTCTGTCAACACAGCAGTCTTTCGGGTGGAACAGATAT 312

DB 240 CTGTCAGCAGACCGGCTTACCTCTGTGAGCAGCAGTCTCTGCGGTGTGAGGCCGAGTAC 299

QY 313 CTCAGTCCGCTTGATTAAGTCAAGTGGCAACAGAACTCCCAAGGAAAGAAATTCGAAG 372

DB 300 CTAGTCCCTTGATTAAGTCTGGAGCCACAGAGCCAGCATTCGAGAGAGTGGCGNA 359

QY 373 GA---AATGAATTTAGCTGTAGGTATGTGGGAGACATTTAGAGTCGCTTTTGAATGT 429

DB 360 GACCCGAGGAGTTGAGCTGTGTGTGGGAGACATTTCCAGTGGCTTTTGTATGTT 419

QY 430 GAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGTAAATGTGCGGAGA 489

DB 420 GAGAGCCACATGAAGAACATGAGTCTTTCAGTATGGGTGAGATGTGCGGGAGG 479

QY 490 AGMTTSSRAGGCTTGGTTTCTTAAAAATCACATGCGGACACATAATGCGAAATCGGG 549

DB 480 AGATTCAGGAGCGCTGCTTCTTGAGAACACATGCGGACACACATGCGCAAGTCTGGC 539

QY 550 GCCAGAGCAAACTGCGACAAAGGCTTGGAGAGTAGTCCAGCAACCATCAAGAGTCTGTC 609

DB 540 ACCAGGAGCAAGCTTCAGCAAGGAGTGGAG---AGTCCAGTCACCATCAATGAAGTGTG 596

QY 610 CAGGTGACGCGGCGGAGAGCATCTCTCTCTTACAAATCTGCATGTTGTGCTTC 669

DB 597 CAGCGCACGCGGCTTGGGAGCATCTCCACGCGGCTTACAAAGATCTGCATGTTGTGCGGCTTC 656

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGTGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

| ORIGIN | Query Match | Best Local Similarity | 14.2%; Score 453.2; DB 6; Length 805; | 78.9%; Pred. No. 1.5e-119; | Matches 633; Conservative 0; Mismatches 145; Indels 24; Gaps 7; |
|--------|-------------|---|---------------------------------------|----------------------------|---|
| QY | 1030 | CTCTCGCAAGAGAAAGTGCAAACTCCACGCGGAGAGCGCCCTCGTGGAGCGG | 1089 | | |
| DB | 1 | CTCTCCCAAGACAAAGAGAGCCCTAGACATGCTAAATAGTAGAGTGCCTCTTGGGATAGT | 60 | | |
| QY | 1090 | GATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTCGGCGCAAGCT | 1149 | | |
| DB | 61 | GACCCCAAGTTGTCAGTAGCAAGGAGAGCCCAACGCACTGTTCTGAGTGCAGCAAGACC | 120 | | |
| QY | 1150 | TTCAGAACCTTACCACACAGCTGGTCTTGCACTCCAGGGTCCACAAGAGGACCGGAGGCC | 1209 | | |
| DB | 121 | TTCAGGACATACACACNGCTCGTCTGCATCTGAGGGTGCACAGGAAGGACAGGAGACT | 180 | | |
| QY | 1210 | GGCGGGAGTCGCCACCATGTCTGTGGACGGGAGCGACGGGGACGTTTCTCTGTAC | 1269 | | |
| DB | 181 | GATGCCCTGTGACCCACCATGGCTGTGGATGCAAGGCAGCTGGGACCTGCTCCCCAGAC | 240 | | |
| QY | 1270 | CTCGCGCCCTCTGGATGAAAATGGAGCCGCTGGATCGAGGGGAGTGTTCTCTGAAGAC | 1329 | | |
| DB | 241 | CTCAGCACCACTCTGGAAGACAGTGGGGCGGGGACCGA--GAAGGGGGCTCTGAAGAC | 297 | | |
| QY | 1330 | GGATCTGAGATGGGCTTCCCGAAGGAATCCATCTCGGATAAAATGATGATGAGGAGAAA | 1389 | | |
| DB | 298 | GGGTCTGAGGATGGACCTCCCTGACGGGCTCATTTGGATGATAAATGATGATGAGGAAAA | 357 | | |
| QY | 1390 | ATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGTGGAAAGTTTTTCCGTTCAAAAT | 1449 | | |
| DB | 358 | GCGAAGCCCTCCGCTCCTCGAGAGAGTGTAGTTACTGTGGCAAGTTTTTCCGTTCAAAAC | 417 | | |
| QY | 1450 | TATTACTCAATATTCATCTCAGAACGCATACAGTGCAGAAAACCATACAAATGTGAATTT | 1509 | | |
| DB | 418 | TATTACTCAATATTCATCTCAGAACGCATACAGTGTGAAAACCATACAAATGTGAATTC | 477 | | |
| QY | 1510 | TGTGAATATGCTGCAGGCCAGAGACATCTCGAGGTATCAC--TTGGAGAGACATACAA | 1568 | | |
| DB | 478 | TGTGAGTATGCCGAGCCGACAGAGACATCTCTGAGGTACCCCTTTGGAGAGACATACAA | 537 | | |
| QY | 1569 | GGAAAAACAAAC--CGATGTTGCTGCTGAAGTCAAGAACGATGTTAAAAATCAGGACAC | 1625 | | |
| DB | 538 | AGACAAAGCAGCCGCTGGATGCTCGCGCTGAGTCCAAAAGTGAAGCCGCGAGCCAGAGCC | 597 | | |
| QY | 1626 | TGAAGATGCACTATTAA---CCGCTGACAGTGGCGCAAAACCAAAAATTTGAAAAAGATTTTT | 1682 | | |
| DB | 598 | GCAAGATGCGCTACTAACCGCTGCTGACAGTGGCGAGACCANAATAATTTAAAGAGATTTCT | 657 | | |
| QY | 1683 | TGATGTGCCAAGATGTTACAGGCAGTGCACCTGCMAAGCAGCTTAAGAGATGCGCTTC | 1742 | | |
| DB | 658 | TGATGTGCCAAGATGTTTA-AGGGAAGCCACCTGCCAAGCAGCTTAAGAGATG-CTTC | 715 | | |
| QY | 1743 | TGTTTTTTCAGAAATGTTCTGGGCGAGCGTGTCTCTCACACGACACACAAAGATACTCAGGA | 1802 | | |
| DB | 716 | TGTCCTCCAGAGTGTTCT-----TCTCNACGACACACAGCAACGATACTCAGGA | 763 | | |
| QY | 1803 | TTTTCCATAAAATGCGAGCTGAT | 1824 | | |
| DB | 764 | TTTTCCATAAACATGCGAGCTGAT | 785 | | |

RESULT 11
BB617224
LOCUS

| LOCUS DEFINITION | ACCESSION VERSION | KEYWORDS SOURCE | ORGANISM |
|---------------------|----------------------|--------------------|----------|
|---------------------|----------------------|--------------------|----------|

REFERENCE AUTHORS

| TITLE | JOURNAL | COMMENT |
|---|-----------------------------------|----------------------------------|
| 1. The Role of the Teacher in the Classroom | Journal of Educational Research | 1980, Vol. 83, No. 1, pp. 1-10 |
| 2. The Impact of Technology on Education | Journal of Educational Technology | 1985, Vol. 10, No. 2, pp. 1-10 |
| 3. The Importance of Parental Involvement | Journal of Educational Research | 1990, Vol. 93, No. 3, pp. 1-10 |
| 4. The Effect of Teacher Education on Student Achievement | Journal of Educational Research | 1995, Vol. 98, No. 4, pp. 1-10 |
| 5. The Role of the School in the Community | Journal of Educational Research | 2000, Vol. 103, No. 5, pp. 1-10 |
| 6. The Impact of Teacher Salary on Student Achievement | Journal of Educational Research | 2005, Vol. 108, No. 6, pp. 1-10 |
| 7. The Role of the School in the Community | Journal of Educational Research | 2010, Vol. 113, No. 7, pp. 1-10 |
| 8. The Impact of Teacher Salary on Student Achievement | Journal of Educational Research | 2015, Vol. 118, No. 8, pp. 1-10 |
| 9. The Role of the School in the Community | Journal of Educational Research | 2020, Vol. 123, No. 9, pp. 1-10 |
| 10. The Impact of Teacher Salary on Student Achievement | Journal of Educational Research | 2025, Vol. 128, No. 10, pp. 1-10 |

BB617224 632 bp cDNA clone 4933431C08 5', mRNA sequence.
BB617224 BB617224.1 GI:15396288
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 692)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
Ohno, M., Okazaki, Y., Okido, T., Saito, K., Sakai, C., Sakai, K.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, P.,
Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y.
Riken Mouse ESTs (Arakawa, T. et al., 2001)

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsic.riken.jp. URL: <http://genome.gsc.riken.jp/>
Carinci, P., Shbata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Ito, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000).
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.,
and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001).

Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T., Ishii, Y. and Hayashizaki, Y.

Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct. Func. Genomics* 2 pre, L72-L86 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
SOURCE

```

Location/Qualifiers
1. .692
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4933431C08"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="SOLR"

```

```
/lib_nost="SOLK"  
/clone_lib="RIKEN full-length enriched, adult male testis"  
/note="Site 1: XhoI; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in
```



```

Db      421  CCGCCCCCTCTGGATGATACATGAGCCGCTGGATCGAGGGAAGGCTGGTCTGAATAAG 478

RESULT 13
W05407
LOCUS      za83c12.r1 Soares fetal lung NBHL19W Homo sapiens cDNA clone
DEFINITION IMAGE:299158 5' similar to SW:2F26_MOUSE P10076 ZINC FINGER PROTEIN
ZFP-26 ;, mRNA sequence.
ACCESSION W05407
VERSION   W05407.1 GI:1278138
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 473)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.
TITLE     The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Contact: Wilson RK
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: est@wustl.wustl.edu
          This clone is available royalty-free through LNL; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          Seq primer: ETPRimer
          High quality sequence stop: 406.
          Location/Qualifiers
            1..473
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="GDB:1244082"
              /db_xref="IMAGE:299158"
              /clone="IMAGE:299158"
              /dev_stage="19 weeks"
              /lab_host="DH10B (ampicillin resistant)"
              /clone_lib="Soares_fetal_lung_NBHL19W"
              /note="Organ: lung; Vector: pT73D (Pharmacia) with a
              modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
              strand cDNA was primed with a Not I - oligo(dT) primer
              [5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTTTTTTTTT-3'],
              double-stranded cDNA was size selected, ligated to Eco RI
              adapters (Pharmacia), digested with Not I and cloned into
              the Not I and Eco RI sites of a modified pT73 vector
              (Pharmacia). Library went through one round of
              normalization to a Cot = 5. Library constructed by Bento
              Soares and M. Fatima Ronaldo. This library was constructed
              from the same fetus as the fetal heart library, Soares
              fetal heart NBHL19W."

FEATURES
source
    Query Match      12.4%; Score 395.2; DB 7; Length 473;
    Best Local Similarity 92.9%; Pred. No. 9e-103;
    Matches 446; Conservative 0; Mismatches 27; Indels 7; Gaps 3;

QY      1106  GTAGCAAGGAGAGCCCACTACTCTCGATGCGGCAAAAGCTTTCAGAACCTACCAACC 1165
Db      1      GTAGCAAGGAGAGCCCACTACTCTCGATGCGGCAAAAGCTTTCAGAACCTACCAACC 60

QY      1166  AGCTGCTCTGCACCTCAGGCTCCACAGAGGACCGGCGCGGAGTGCCTCCCA 1225
Db      61      AGCTGCTCTGCACCTCAGGCTCCACAGAGGACCGGCGCGGAGTGCCTCCCA 118

QY      1226  CCATGCTGTGACGGGAGGACCGGAGCGTGTCTCTGACCTCGCCGCCCTCTGG 1285
Db      119     CCATGCTGTGACGGGAGGACCGGAGCGTGTCTCTGACCTCGCCGCCCTCTGG 178

```

```

QY      1286  ATGAAATGGAGCCGCTGATCGAGGGGAAGTGGTCTTGAAGACCGATCTGAGGATGGC 1345
Db      179     ATGAAATGGAGCCGCTGATCGAGGGGAAGTGGTCTTGAAGACCGATCTGAGGATGGC 238

QY      1346  TTCCCGAAGGAATCCATCTCGATAAAAAATGATGAGGAGGAAAAATAAAACATCTTACAT 1405
Db      239     TTCCCGATAGATCCATCTCGATAAAAAATGATGAGGAGGAAAAATAAAACATCTTACAT 298

QY      1406  CTTCAAGAGAGTGTAGTTATTGTGAAAGTTTTTCCGTTCAAAATTTTACCTCAATATTC 1465
Db      299     CTTCAAGAGAGTGTAGTTATTGTGAAAGTTTTTCCGTTCAAAATTTTACCTCAATATTC 358

QY      1466  ATCTCAGAACCCATACAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGAG 1525
Db      359     ATCTCAGAACCCATACAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGAG 418

QY      1526  CCCAGNAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAAGGAAAAACAAACCGATG 1585
Db      419     CCAGNAGACATCT-----GAGGTACACTTGGAGAGCAACAGGAAANNAACAACCGATG 473

RESULT 14
BG007233
LOCUS      IL5-GN0239-271100-281-b03 GN0239 Homo sapiens cDNA, mRNA sequence.
DEFINITION BG007233
ACCESSION  BG007233
VERSION    BG007233.1 GI:12451213
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 411)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE   20202663
PUBMED    10737800
COMMENT   Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=IL5&t2=IL5-GN0239-
          271100-281-b03&t3=2000-11-27&t4=1)
          Seq primer: puc 18 forward
          High quality sequence stop: 360.
          Location/Qualifiers
            1..411
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /dev_stage="Adult"
              /clone_lib="GN0239"
              /note="Organ: placenta_normal; Vector: puc18; Site 1:
              SmaI; Site 2: SmaI; A mini-library was made by cloning
              products derived from ORESTES PCR (U.S. Letters Patent
              application No. 196,716 - Ludwig Institute for Cancer
              Research) profiles into the puc 18 vector. Reverse
              transcription of tissue mRNA and cDNA amplification were
              performed under low stringency conditions."

FEATURES
source
    Query Match      12.4%; Score 395.2; DB 7; Length 473;
    Best Local Similarity 92.9%; Pred. No. 9e-103;
    Matches 446; Conservative 0; Mismatches 27; Indels 7; Gaps 3;

QY      1106  GTAGCAAGGAGAGCCCACTACTCTCGATGCGGCAAAAGCTTTCAGAACCTACCAACC 1165
Db      1      GTAGCAAGGAGAGCCCACTACTCTCGATGCGGCAAAAGCTTTCAGAACCTACCAACC 60

QY      1166  AGCTGCTCTGCACCTCAGGCTCCACAGAGGACCGGCGCGGAGTGCCTCCCA 1225
Db      61      AGCTGCTCTGCACCTCAGGCTCCACAGAGGACCGGCGCGGAGTGCCTCCCA 118

QY      1226  CCATGCTGTGACGGGAGGACCGGAGCGTGTCTCTGACCTCGCCGCCCTCTGG 1285
Db      119     CCATGCTGTGACGGGAGGACCGGAGCGTGTCTCTGACCTCGCCGCCCTCTGG 178

```

ORIGIN

```
Query Match      12.4%; Score 394; DB 4; Length 411;
Best Local Similarity 97.3%; Pred. No. 1.9e-102; Indels 0; Gaps 0;
Matches 400; Conservative 0; Mismatches 11;

Qy 2433 CACTTTAGCCCCAAGTAACCTGAAGTCCACAGACCACAGAGAATGTGGGGTCCAAGG 2492
Db 1 CACGAGCGCTCCAGTAACCTGAAGTCCACAGACCACAGAGAATGTGGAGATCCAAGG 60

Qy 2493 GCGCCGACACAGGCAACGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCGTCGACC 2552
Db 61 GCGCCGACACAGGCAACGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCGTCGACC 120

Qy 2553 GGATAAGACAAAAGACCCGAGACAAAATGAAACCTCTTCAGTAGCTCTTCTTCAGCC 2612
Db 121 GGATAAGACAAATAGACCCGAGACAAAATGAAACCTCTTCAGTAGCTCTTCTTCAGCC 180

Qy 2613 CACCCTCGGACGAGTAACATCAATGGTTCCATCGACTACCCCGCAAGAACGACAGCCC 2672
Db 181 CACCCTCGGACGAGTAACATCAATGGTTCCATCGACTACCCCGCAAGAACGACAGCCC 240

Qy 2673 GTGGCCACCTCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATACCTGACGAGA 2732
Db 241 GTGGCCACCTCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATACCTGACGAGA 300

Qy 2733 ATTTGGTCAGCCCTTCCAAAAGACTCAAGTCCAGCGTGGTTGCGCTTGACGTTGACCA 2792
Db 301 ATTTGGTCAGCCCTTCCAAAAGACTCAAGTCCAGCGTGGTTGCGCTTGACGTTGACCA 360

Qy 2793 GCGCGGGCCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGT 2843
Db 361 GCGCGGNGCCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGT 411
```

RESULT 15

```
BM834501
LOCUS K-BST0109502 S11SNUI Homo sapiens cDNA clone S11SNUI-63-D04 5',
        mRNA sequence.
ACCESSION BM834501
VERSION BM834501.1 GI:19190910
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
```

21C Frontier Korean EST Project 2001

TITLE

Unpublished (2002)

JOURNAL

Genome Research Center

COMMENT

Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 63 row: D column: 04
High quality sequence stop: 640.

FEATURES

source

```
1..640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S11SNUI-63-D04"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/lab_host="SNU-1"
/lab_host="Top10F"
```

/clone_lib="S11SNUI"

/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including Sfil
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using Sfil
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with Sfil and
cloned into Bralll- digested pME18S-FL3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

```
Query Match      12.1%; Score 384; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 1.8e-99;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCAATCGAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGG 60
Db 257 ATGCAATCGAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGG 316

Qy 61 CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 120
Db 317 CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 376

Qy 121 AAAGGGACCGCTGTTGTTCCATCCGAGCTACACAAGAAAAAATGTGCATCCAAATCGAG 180
Db 377 AAAGGGACCGCTGTTGTTCCATCCGAGCTACACAAGAAAAAATGTGCATCCAAATCGAG 436

Qy 181 GGGTATATGCCCTTGGATTGCAATGTTCTGAGCCAGACCTTCACACATTCAGAGACCTT 240
Db 437 GGGTATATGCCCTTGGATTGCAATGTTCTGAGCCAGACCTTCACACATTCAGAGACCTT 496

Qy 241 AATAAACATGTCTTAATGCAACACCGGCTACCTCTGTGAACAGCAGATGTTCTTCGGGTT 300
Db 497 AATAAACATGTCTTAATGCAACACCGGCTACCTCTGTGAACAGCAGATGTTCTTCGGGTT 556

Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCAGACAGAACCTCCCAAGGAA 360
Db 557 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCAGACAGAACCTCCCAAGGAA 616

Qy 361 AAGAAATTCAGAGGAAATGAATTT 384
Db 617 AAGAAATTCAGAGGAAATGAATTT 640
```

Search completed: July 3, 2005, 16:40:02

Job time : 10081.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:48:07 ; Search time 13000 Seconds
(without alignments)
11428.001 Million cell updates/sec

Title: US-08-731-499-12
Perfect score: 3066
Sequence: 1 GGAACAGCTATGACCATGA.....GGTACCCTATCGCCCTATA 3066

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 3066 | 100.0 | 3066 | 6 | BD085735 Genes fro |
| 2 | 2813 | 91.7 | 2813 | 6 | CQ776693 Sequence |
| 3 | 2813 | 91.7 | 2813 | 6 | AR371889 Sequence |
| 4 | 2813 | 91.7 | 2813 | 9 | AF041260 Homo sapi |
| 5 | 2740.4 | 89.4 | 2808 | 6 | CQ726468 Sequence |
| 6 | 1016.4 | 33.2 | 163196 | 9 | AC005220 Homo sapi |
| 7 | 912.8 | 29.8 | 2916 | 10 | AY219233 |
| 8 | 584.6 | 19.1 | 39507 | 9 | AC004501 |
| 9 | 583 | 19.0 | 181313 | 2 | AC021963 |
| 10 | 470.6 | 15.3 | 1885 | 10 | AB029897 |
| 11 | 470.4 | 15.3 | 488 | 6 | AX150111 |
| 12 | 303.6 | 9.9 | 105616 | 10 | AL935134 |
| 13 | 303.6 | 9.9 | 168876 | 2 | AC126535 |
| 14 | 303.6 | 9.9 | 233233 | 2 | AC094145 |
| 15 | 303.6 | 9.9 | 235411 | 2 | AC084066 |
| 16 | 265 | 8.6 | 265 | 6 | CQ112058 Sequence |
| 17 | 265 | 8.6 | 265 | 6 | CQ150869 Sequence |
| 18 | 265 | 8.6 | 265 | 6 | CQ234209 Sequence |
| 19 | 265 | 8.6 | 265 | 6 | CQ309538 Sequence |

| | | | | | | | |
|----|-------|-------|--------|--------|----------|-----------|-----------|
| 20 | 265 | 8.6 | 265 | 6 | CQ346160 | Sequence | |
| 21 | 264.2 | 8.6 | 724 | 10 | BC048647 | Mus muscu | |
| 22 | 260.4 | 8.5 | 495 | 6 | CQ099004 | Sequence | |
| 23 | 260.4 | 8.5 | 495 | 6 | CQ137928 | Sequence | |
| 24 | 260.4 | 8.5 | 495 | 6 | CQ221300 | Sequence | |
| 25 | 260.4 | 8.5 | 495 | 6 | CQ297119 | Sequence | |
| 26 | 260.4 | 8.5 | 495 | 6 | CQ333492 | Sequence | |
| 27 | 247 | 8.1 | 349 | 6 | BD060529 | Secreted | |
| c | 284.4 | 8.0 | 106018 | 9 | HS864118 | Human DNA | |
| 29 | 241.2 | 7.9 | 228652 | 9 | CNS01DWD | Human chr | |
| 30 | 240.8 | 7.9 | 180914 | 9 | CNS01DS1 | Human chr | |
| c | 31 | 240.8 | 7.9 | 181183 | 2 | AL391866 | Homo sapi |
| 32 | 240.8 | 7.9 | 214527 | 9 | CNS0000J | Human chr | |
| c | 33 | 240.6 | 7.8 | 686 | 11 | AF003745 | Human chr |
| c | 34 | 238.4 | 7.8 | 138573 | 9 | AL954211 | Pan trogl |
| 35 | 238.2 | 7.8 | 105208 | 9 | AL355515 | Human DNA | |
| 36 | 238 | 7.8 | 297 | 6 | AX898536 | Sequence | |
| 37 | 238 | 7.8 | 297 | 6 | BD034069 | Sequence | |
| 38 | 237.6 | 7.7 | 172510 | 9 | AL512424 | Human DNA | |
| 39 | 237.4 | 7.7 | 159667 | 9 | AC078953 | Homo sapi | |
| c | 40 | 237.2 | 7.7 | 173804 | 9 | HS1030M6 | Human DNA |
| 41 | 236.8 | 7.7 | 144189 | 9 | HS417M14 | Human DNA | |
| c | 42 | 236.8 | 7.7 | 152322 | 2 | AL357077 | Homo sapi |
| c | 43 | 236.8 | 7.7 | 196662 | 9 | AF001107 | Homo sapi |
| c | 44 | 236.6 | 7.7 | 149253 | 2 | AC012245 | Homo sapi |
| 45 | 236.6 | 7.7 | 176132 | 2 | AC061976 | Homo sapi | |

ALIGNMENTS

RESULT 1
BD085735
LOCUS
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085735
VERSION BD085735.1 GI:22631345
KEYWORDS JP 2001524802-A/11.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 3066)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 11 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
COMMENT OS Artificial Sequence
PN JP 2001524802-A/11
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11, C12Q1/68, A61K48/00
CC Description of Artificial Sequence:lb1
FH Key Location/Qualifiers
FT source 1..3066
FT /organism='Artificial Sequence'.
FEATURES
source
1..3066
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 100.0%; Score 3066; DB 6; Length 3066;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3066; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGAACAGCTATGACCATGATTAGCCCAAGCTCGAATTAACCTCACTAAGGGAACAA 60

Db 1 GGAACAGCTATGACCATGATTAGCCCAAGCTCGAAATTAACCCCTCACTAAGGGAACAA 60
Qy 61 AAGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGTATCCCCGGGCTCGAG 120
Db 61 AAGCTGGAGCTCCACCGCGTGGCGCGCTCTAGAACTAGTGTATCCCCGGGCTCGAG 120
Qy 121 GAAATTCGGCAGCGGCTCCACCGCAGACCGAGGCACTGGGCGACGACGACTGGAGACCCAG 180
Db 121 GAAATTCGGCAGCGGCTCCACCGCAGACCGAGGCACTGGGCGACGACGACTGGAGACCCAG 180
Qy 181 GACCCCTGTGACGAGCAGCTCCGGGTGACACGAGGGGACTGGAAGATACTCCCAAGGGGC 240
Db 181 GACCCCTGTGACGAGCAGCTCCGGGTGACACGAGGGGACTGGAAGATACTCCCAAGGGGC 240
Qy 241 TCAGCAGGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGAT 300
Db 241 TCAGCAGGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGAT 300
Qy 301 GAACAGAGCAGAGACTTACAGGACAAACCGCTCTGCTCTGAACGGGGTTCCAGTGGTG 360
Db 301 GAACAGAGCAGAGACTTACAGGACAAACCGCTCTGCTCTGAACGGGGTTCCAGTGGTG 360
Qy 361 GTGTGACCCACACAGTTTACGACTTAGAGAAAGTTCGACTTGGGAATAAGTGTCAAGACG 420
Db 361 GTGTGACCCACACAGTTTACGACTTAGAGAAAGTTCGACTTGGGAATAAGTGTCAAGACG 420
Qy 421 GATAATGTGGCCACTTCTTCCCGGAGACACGAGATAGTGTGTGGGATGCCAAC 480
Db 421 GATAATGTGGCCACTTCTTCCCGGAGACACGAGATAGTGTGTGGGATGCCAAC 480
Qy 481 GGAAGAAATCTTGGGAAGAGGCCAAACCCGAGGACCCAGCTGTCTAAATCTCGTTTTTTC 540
Db 481 GGAAGAAATCTTGGGAAGAGGCCAAACCCGAGGACCCAGCTGTCTAAATCTCGTTTTTTC 540
Qy 541 TTGATGTCTCTCGGCTGTACAGGAGCTTACCGGAGACGTAACCGGAGACCAAGCGGAGATTCATCCCTT 600
Db 541 TTGATGTCTCTCGGCTGTACAGGAGCTTACCGGAGACGTAACCGGAGACCAAGCGGAGATTCATCCCTT 600
Qy 601 GGATCAGTGAAGCTTGTAGTGTAGTCCCAATTAAGTCCAGGCAACAGACCCCAAGTGAG 660
Db 601 GGATCAGTGAAGCTTGTAGTGTAGTCCCAATTAAGTCCAGGCAACAGACCCCAAGTGAG 660
Qy 661 AGCTGGACACTTCGCGTGGAGCTGGACCGGGGAGGACACAGATAAAACCCCAAGGGCAC 720
Db 661 AGCTGGACACTTCGCGTGGAGCTGGACCGGGGAGGACACAGATAAAACCCCAAGGGCAC 720
Qy 721 GCGCGGCCCCAAGACAAGGTCTCTCTGCGCGCAGGGATCCACGCTTCTCCCACTGAG 780
Db 721 GCGCGGCCCCAAGACAAGGTCTCTCTGCGCGCAGGGATCCACGCTTCTCCCACTGAG 780
Qy 781 ACAGGGGAGCAGGAGAGAGCTCCCTCCAGGCCCAAGGACTCCAGCTTTTGTGACAA 840
Db 781 ACAGGGGAGCAGGAGAGAGAGCTCCCTCCAGGCCCAAGGACTCCAGCTTTTGTGACAA 840
Qy 841 TTCTTCAAGCTGGACAGGAGACAGGAAAGTGCAGGTGACAGCAACAGAGAGCCCAAG 900
Db 841 TTCTTCAAGCTGGACAGGAGACAGGAAAGTGCAGGTGACAGCAACAGAGAGCCCAAG 900
Qy 901 AGGGCAGAGCATCAAGACAAGGTGGATGAGTTCCTGGCTTATCAGGCGAGTCCGATGAT 960
Db 901 AGGGCAGAGCATCAAGACAAGGTGGATGAGTTCCTGGCTTATCAGGCGAGTCCGATGAT 960
Qy 961 GTCCCTGCGGAGGACATAGTTGACCGGCAAGGAAAGAGGACAAAGAACTTGGAACT 1020
Db 961 GTCCCTGCGGAGGACATAGTTGACCGGCAAGGAAAGAGGACAAAGAACTTGGAACT 1020
Qy 1021 GCGGATTCCTCTGCTCCCTGGGACCCAGAGGACTGGAGACTGCAAGAGGAGATTCACG 1080
Db 1021 GCGGATTCCTCTGCTCCCTGGGACCCAGAGGACTGGAGACTGCAAGAGGAGATTCACG 1080
Qy 1081 GCAGCAGCTATAGCAGAGAAATAATTCATGAGTTCTTTAAAACTCTGGTTTCA 1140

Db 1081 GCAGCAGCTATAGCAGAGAAATAATTCATGAGTTTCTTTAAAACTCTGGTTTCA 1140
Qy 1141 CCTAAACAAAGCTGAAACAAAAAGGACCCAGAGACACCGGTGCTGAAAAAGTCAACCAAC 1200
Db 1141 CCTAAACAAAGCTGAAACAAAAAGGACCCAGAGACACCGGTGCTGAAAAAGTCAACCAAC 1200
Qy 1201 ACTTCAGCTGACCTTAAAGTCCAGACAAAGCCAACTTTACATCCCAAGGAGACCAAGGGCT 1260
Db 1201 ACTTCAGCTGACCTTAAAGTCCAGACAAAGCCAACTTTACATCCCAAGGAGACCAAGGGCT 1260
Qy 1261 GGCAGAAATTCAAAGAGTGCACCCATCGGGGACACACAGTCCGTCGACAAACCCCTGAA 1320
Db 1261 GGCAGAAATTCAAAGAGTGCACCCATCGGGGACACACAGTCCGTCGACAAACCCCTGAA 1320
Qy 1321 CCTGCGAAGAGCACCACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
Db 1321 CCTGCGAAGAGCACCACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
Qy 1381 TTTTGGAAAAAGTCAAGTAAAGAGGACTCAGTCCCAAGTTCGCGGAGGAGAAATGTGGTG 1440
Db 1381 TTTTGGAAAAAGTCAAGTAAAGAGGACTCAGTCCCAAGTTCGCGGAGGAGAAATGTGGTG 1440
Qy 1441 TGTGAGTCAACGAGTAGAGATTATAAGTCCAAGGAAGTAGAATCAGCTTACAAACAGTG 1500
Db 1441 TGTGAGTCAACGAGTAGAGATTATAAGTCCAAGGAAGTAGAATCAGCTTACAAACAGTG 1500
Qy 1501 GACCTCAACGAAGGAGATGCTGCACCTGAAACCCACAGAGGAACTCAAAAGAGAGAA 1560
Db 1501 GACCTCAACGAAGGAGATGCTGCACCTGAAACCCACAGAGGAACTCAAAAGAGAGAA 1560
Qy 1561 AGCAAAACCAAGAACTCTCTGATGGCGTTTCTCAGACAAATGTCAAGTAAAGGGATGGA 1620
Db 1561 AGCAAAACCAAGAACTCTCTGATGGCGTTTCTCAGACAAATGTCAAGTAAAGGGATGGA 1620
Qy 1621 GGGATCAACCACTCAGAAAGAAATAATGGGAAGACTCCAGCTCCCAACATCAGACTCC 1680
Db 1621 GGGATCAACCACTCAGAAAGAAATAATGGGAAGACTCCAGCTCCCAACATCAGACTCC 1680
Qy 1681 ACAGAAAGACTATCACACCGCAGAGCTGAAACCAACAGGAGACCAACAGAGGGTAAA 1740
Db 1681 ACAGAAAGACTATCACACCGCAGAGCTGAAACCAACAGGAGACCAACAGAGGGTAAA 1740
Qy 1741 GAGGGCTCTCGAAGGACAAAGAGTCAAGAGTCCAGAGCCGAGATGAACAGAGAGAGCAACAG 1800
Db 1741 GAGGGCTCTCGAAGGACAAAGAGTCAAGAGTCCAGAGCCGAGATGAACAGAGAGAGCAACAG 1800
Qy 1801 CAGAAACCAAGAGACCAAGCCAGTGCACAGAGAGCCGCGGTGGACACCACTCACTG 1860
Db 1801 CAGAAACCAAGAGACCAAGCCAGTGCACAGAGAGCCGCGGTGGACACCACTCACTG 1860
Qy 1861 CAGAAATGGGACAAAGCTCCAAAGAGACCTGAGAAAGCGGACAGTCCCTTGGGGGCTTC 1920
Db 1861 CAGAAATGGGACAAAGCTCCAAAGAGACCTGAGAAAGCGGACAGTCCCTTGGGGGCTTC 1920
Qy 1921 TTTAAAGCGCTGGACCAAGCGGATGTGGATGCTCAAGTGCAACAGAGACCAAGTATCC 1980
Db 1921 TTTAAAGCGCTGGACCAAGCGGATGTGGATGCTCAAGTGCAACAGAGACCAAGTATCC 1980
Qy 1981 ATCGGACAGTGTGGCAACCCCAAGTAAACAAATCAGCAGCGTTCCCAAGGTTCTCCTG 2040
Db 1981 ATCGGACAGTGTGGCAACCCCAAGTAAACAAATCAGCAGCGTTCCCAAGGTTCTCCTG 2040
Qy 2041 CCACAAAGATGTGTCTCTTACT 2100
Db 2041 CCACAAAGATGTGTCT 2100
Qy 2101 TTCTGATGGCCAGCAATGAAATTTCTGCTTAGAAATTTAAGCCCGAGCTGTGTATATGA 2160
Db 2101 TTCTGATGGCCAGCAATGAAATTTCTGCTTAGAAATTTAAGCCCGAGCTGTGTATATGA 2160
Qy 2161 GGTGTATATTATGCTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
Db 2161 GGTGTATATTATGCTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220

| | | | |
|----|------|---|------|
| QY | 2221 | GCAGGTACCTAGTGGGTGACAGAGTCGATGATCACCACGAGGAAGGAGGGAATA | 2280 |
| Db | 2221 | | |
| QY | 2281 | GAGGAATGTGTTGCGGTAAAGTGATGAAATGGCAGTGTGCGCGGGGTGCTCTC | 2340 |
| Db | 2281 | | |
| QY | 2341 | GCCTGTATCTCAGCACTTTGGGAGCGGAGGAGGTGATCACCCTGAGGTGAGAGTTC | 2400 |
| Db | 2341 | | |
| QY | 2401 | AAGACTAGCTGCGCCCAACATCATGAAACCCCGTCTCTAATAAATACAAAATTTAGCCA | 2460 |
| Db | 2401 | | |
| QY | 2461 | GGCATGTGGGCACACACCTGTAGTCCCACTACTCGGAGGCCAACCGACGAAACCGCT | 2520 |
| Db | 2461 | | |
| QY | 2521 | TGTACCCAGGAGGTGGAGTTGCACTGAGCGGAGTTCGACCATTCACCTCCGCTGG | 2580 |
| Db | 2521 | | |
| QY | 2581 | CGACAGCAGATTTCTATCAAAAAAAGAGCAGTGGCAAGTAAGTTATAGAACAGAA | 2640 |
| Db | 2581 | | |
| QY | 2641 | TGCTCTAGAGGAATTAAGCGTGTAGTAAACGGTGTCTCATCTCTAAAGCTTCAAGAA | 2700 |
| Db | 2641 | | |
| QY | 2701 | GGGAGACGAAATCCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCCGGCTGTGT | 2760 |
| Db | 2701 | | |
| QY | 2761 | TGGGTGTTGCCAATTTCTAGAACGGAATGTGGGTATAGAAAAAGGATGAATAAG | 2820 |
| Db | 2761 | | |
| QY | 2821 | CGTTGTTTTCAATAGGTCCTTGTAAAGTTATTGATGAGGGGAAAAAGATTGACTGGG | 2880 |
| Db | 2821 | | |
| QY | 2881 | AGGGCTTAAATGATTTGGGAAAACAATTTGCTTTTGTAGGCTCAGTGAACAACGGCAAGAT | 2940 |
| Db | 2881 | | |
| QY | 2941 | TACAACTTAAAAAATAAAAAAATAAAAACTCGAGACTAGTTCTCTCTCTCTCGTGCCGA | 3000 |
| Db | 2941 | | |
| QY | 3001 | ATTGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCGGTACCCTTCCG | 3060 |
| Db | 3001 | | |
| QY | 3061 | CCTATA 3066 | |
| Db | 3061 | | |

RESULT 2
CQ776693
LOCUS CQ776693 2813 bp DNA linear PAT 11-MAR-2004
DEFINITION Sequence 379 from Patent EP1394274.
ACCESSION CQ776693
VERSION CQ776693.1 GI:45380083
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

| | | |
|-----------------------|----------------------------|--|
| REFERENCE | 1 | Ohtani, N., Sugita, Y., Yamaya, M., Kubo, H., Nagai, H. and Izuwara, K. Methods of testing for bronchial asthma or chronic obstructive pulmonary disease Patent: EP 1394274-A 379 03-MAR-2004; Genex Research, Inc. (JP) |
| AUTHORS | | |
| JOURNAL | | |
| FEATURES | | |
| source | 1..2813 | Location/Qualifiers |
| | /organism="Homo sapiens" | |
| | /mol_type="unassigned DNA" | |
| | /db_xref="taxon:9606" | |
| ORIGIN | | |
| Query Match | 91.7%; | Score 2813; DB 6; Length 2813; |
| Best Local Similarity | 100.0%; | Pred. No. 0; |
| Matches 2813; | Conservative 0; | Mismatches 0; Indels 0; Gaps 0; |
| QY | 136 | CTCCACCCACAGCCAGGACCTGGGCGAGCAGCGACCTGGAGACCAGGACCCCTGTGCAGGAG 195 |
| Db | 1 | CTCCACCCACAGCCAGGACCTGGGCGAGCAGCGACCTGGAGACCAGGACCCCTGTGCAGGAG 60 |
| QY | 196 | CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCACAGGGGCTCAGCAGGAGCAATG 255 |
| Db | 61 | CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCACAGGGGCTCAGCAGGAGCAATG 120 |
| QY | 256 | GGTAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGATGAACCAAGAGCAGAG 315 |
| Db | 121 | GGTAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGATGAACCAAGAGCAGAG 180 |
| QY | 316 | ACTTACACGAGCAACCGCTCTGCTGAACCGGGTTCAGTGGTGTGTCGACCCACACA 375 |
| Db | 181 | ACTTACACGAGCAACCGCTCTGCTGAACCGGGTTCAGTGGTGTGTCGACCCACACA 240 |
| QY | 376 | GTTCAGCACTTAGAGGAAAGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 435 |
| Db | 241 | GTTCAGCACTTAGAGGAAAGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 300 |
| QY | 436 | TCTTCCCGGAGACAAACCGGAGATAAGTGTCTGTTCGGATGCCAACGGAAAGATCTTGGG 495 |
| Db | 301 | TCTTCCCGGAGACAAACCGGAGATAAGTGTCTGTTCGGATGCCAACGGAAAGATCTTGGG 360 |
| QY | 496 | AAAGAGGCCAAACCCGAGGCCACGACTCTCTAAATCTCGTTTCTTCTGATGCTCTCTCGG 555 |
| Db | 361 | AAAGAGGCCAAACCCGAGGCCACGACTCTCTAAATCTCGTTTCTTCTGATGCTCTCTCGG 420 |
| QY | 556 | CCTGTACAGGACGTACCGGAGACCAAGCCGAGATTCATCCCTTTGGATCAGTGAAGCTT 615 |
| Db | 421 | CCTGTACAGGACGTACCGGAGACCAAGCCGAGATTCATCCCTTTGGATCAGTGAAGCTT 480 |
| QY | 616 | GATGTACAGTCCAATAAAGTCTCAGCGAAACAAAGACCCCAAGTGAAGCTGGACACTTCGG 675 |
| Db | 481 | GATGTACAGTCCAATAAAGTCTCAGCGAAACAAAGACCCCAAGTGAAGCTGGACACTTCGG 540 |
| QY | 676 | GTGCGAGTGGACCGGGGCGAGACACAGATAAACCAGGAGCCAGCCCGGCCCAAGAC 735 |
| Db | 541 | GTGCGAGTGGACCGGGGCGAGACACAGATAAACCAGGAGCCAGCCCGGCCCAAGAC 600 |
| QY | 736 | AAGGTCTCTCTGCGCGCAGGGATCCCAAGCTTCTTCCACCTTGAGACAGGGGAGCAGGA 795 |
| Db | 601 | AAGGTCTCTCTGCGCGCAGGGATCCCAAGCTTCTTCCACCTTGAGACAGGGGAGCAGGA 660 |
| QY | 796 | GGAGAGCTCCCTCAAGCCCAAGGATTCAGCTTTTGTGACAAATCTTCAAGCTGGAC 855 |
| Db | 661 | GGAGAGCTCCCTCAAGCCCAAGGATTCAGCTTTTGTGACAAATCTTCAAGCTGGAC 720 |
| QY | 856 | AAGGACAGGAAAGGTCAGAGTTCACAGCCAAACAGGAGCCAGGAGGAGAGATCAA 915 |
| Db | 721 | AAGGACAGGAAAGGTCAGAGTTCACAGCCAAACAGGAGCCAGGAGGAGAGATCAA 780 |
| QY | 916 | GACAAGGTGGATGAGGTTCTCTGGCTTATCAGGGCAGTCCGATGATGTCTCTGACGGAGG 975 |
| Db | 781 | GACAAGGTGGATGAGGTTCTCTGGCTTATCAGGGCAGTCCGATGATGTCTCTGACGGAGG 840 |

Qy 976 GACATAGTTGACGCAAGAAAGAAAGGACAAGAACTTGGAACTGGGATGCTCTGTC 1035
Db 841 GACATAGTTGACGCAAGAAAGAAAGGACAAGAACTTGGAACTGGGATGCTCTGTC 900
Qy 1036 CCTGGGACCCAGAAAGGACTGGAGACTGCAAGAGACGATTCGCCAGGACGAGCTATAGCA 1095
Db 901 CCTGGGACCCAGAAAGGACTGGAGACTGCAAGAGACGATTCGCCAGGACGAGCTATAGCA 960
Qy 1096 GAGAAATATAATTCATCATCAGTTTCTTTAAAACTCTGGTTTCACTTAAACAAAGCTGAA 1155
Db 961 GAGAAATATAATTCATCATCAGTTTCTTTAAAACTCTGGTTTCACTTAAACAAAGCTGAA 1020
Qy 1156 ACAAAAAGGACCCAGAAAGACACGGGTGCTGAAAAGTCACCCACCTTCACGTGACCTT 1215
Db 1021 ACAAAAAGGACCCAGAAAGACACGGGTGCTGAAAAGTCACCCACCTTCACGTGACCTT 1080
Qy 1216 AAGTCAGAACCAAGCCAACTTTACATCCAGAGAGACCCAGGGCTGGCAAGATTCNAA 1275
Db 1081 AAGTCAGAACCAAGCCAACTTTACATCCAGAGAGACCCAGGGCTGGCAAGATTCNAA 1140
Qy 1276 GGATGCAACCCATCGGGGACACACAGTCCTGCAACCCCTGAACTTGGCAAGGAGGC 1335
Db 1141 GGATGCAACCCATCGGGGACACACAGTCCTGCAACCCCTGAACTTGGCAAGGAGGC 1200
Qy 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAGAAAGTCA 1395
Db 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAGAAAGTCA 1260
Qy 1396 GTTAAAGAGGACTCAGTCCCAAGGTCGGAGGAGGAAATGTGTGAGTCAACCAGTA 1455
Db 1261 GTTAAAGAGGACTCAGTCCCAAGGTCGGAGGAGGAAATGTGTGAGTCAACCAGTA 1320
Qy 1456 GAGATTATAAAGTCCAAAGGAGTAGAATCAGCTTTACAAACAGTGGACCTCAAACGAAGGA 1515
Db 1321 GAGATTATAAAGTCCAAAGGAGTAGAATCAGCTTTACAAACAGTGGACCTCAAACGAAGGA 1380
Qy 1516 GATGTGCACCTGAACCCACAGAGAGGAACTCAAAGAGAGAAAGCAAAACCAAGAAC 1575
Db 1381 GATGTGCACCTGAACCCACAGAGAGGAACTCAAAGAGAGAAAGCAAAACCAAGAAC 1440
Qy 1576 TCTCTGATGGGTTTTCTCAGACAAATGTTCAGTAAAGGGGATGGAGGATCACCCACTCA 1635
Db 1441 TCTCTGATGGGTTTTCTCAGACAAATGTTCAGTAAAGGGGATGGAGGATCACCCACTCA 1500
Qy 1636 GAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGAGCTATC 1695
Db 1501 GAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGAGCTATC 1560
Qy 1696 ACACGCCAGAGCCTGAACCAACAGGAGCACACAGAGGGTAAAGAGGGCTCCTCGAAG 1755
Db 1561 ACACGCCAGAGCCTGAACCAACAGGAGCACACAGAGGGTAAAGAGGGCTCCTCGAAG 1620
Qy 1756 GACAAGAGTACAGACCGCAGATGAACAGAGAGAGACCAACAGAGAGAGGAGCCAAAGAA 1815
Db 1621 GACAAGAGTACAGACCGCAGATGAACAGAGAGAGACCAACAGAGAGAGGAGCCAAAGAA 1680
Qy 1816 CCAGCCCAAGTGCACAGAGAGCCAGCTGGAGACAGAACTCACTCGAAGTGGGACAAAG 1875
Db 1681 CCAGCCCAAGTGCACAGAGAGCCAGCTGGAGACAGAACTCACTCGAAGTGGGACAAAG 1740
Qy 1876 CTCAAAAAGAGACCTGAGAAGCGGACAGTCCCTTGGGGGCTCTTTAAAGGCTGGGA 1935
Db 1741 CTCAAAAAGAGACCTGAGAAGCGGACAGTCCCTTGGGGGCTCTTTAAAGGCTGGGA 1800
Qy 1936 CCAAAGCGGATGTTGGATGCTCAAGTGCNAAACAGACCCAGTATCATCGACCAAGTTGGC 1995
Db 1801 CCAAAGCGGATGTTGGATGCTCAAGTGCNAAACAGACCCAGTATCATCGACCAAGTTGGC 1860
Qy 1996 AAACCCCAAGTAAACAAATCAGACCGGTTCCACACAGGTTCTCCTGCCACCAAGATGTGT 2055
Db 1861 AAACCCCAAGTAAACAAATCAGACCGGTTCCACACAGGTTCTCCTGCCACCAAGATGTGT 1920
Qy 2056 CTCCTTACTCCATCTCTCCCCAAACACGCTCCATGTATATATTTCTGTGATGGCCAGCA 2115

Db 1921 CTCCTTACTCCATCTCTCCCAACACGCTCCATGTATATATTTCTGTGATGCCAGCA 1980
Qy 2116 AATGAAATTCCTGCTAGAAATTAAGCCGAGCTGTTGTATATTGAGGTGTATTTTACG 2175
Db 1981 AATGAAATTCCTGCTAGAAATTAAGCCGAGCTGTTGTATATTGAGGTGTATTTTACG 2040
Qy 2176 TCTCTGCTCCAGTCTTTTGGCAATAACAGATAAGATGGTTTTAGCAGGTCACCTAGTT 2235
Db 2041 TCTCTGCTCCAGTCTTTTGGCAATAACAGATAAGATGGTTTTAGCAGGTCACCTAGTT 2100
Qy 2236 GGGTGAGAAGAGTCGATGATCAACAGCAGAAAGGAGGGAATAGAGGAATGTTTCGG 2295
Db 2101 GGGTGAGAAGAGTCGATGATCAACAGCAGAAAGGAGGGAATAGAGGAATGTTTCGG 2160
Qy 2296 GTTAAGTGATGAAATGCGAGTGGTGGCCGGCTGTGGCTCTCGCTGTAACTCTCAGC 2355
Db 2161 GTTAAGTGATGAAATGCGAGTGGTGGCCGGCTGTGGCTCTCGCTGTAACTCTCAGC 2220
Qy 2356 ACTTTGGGAGGCCGAGGACGAGTGGATCACCTGAGGTGAGGATTTCAAGACTAGCTGGCC 2415
Db 2221 ACTTTGGGAGGCCGAGGACGAGTGGATCACCTGAGGTGAGGATTTCAAGACTAGCTGGCC 2280
Qy 2416 AACATCATGAAACCCCGTCTCTACTTAAATAACAAAAATTAGCCAGGATGTTGGCACAC 2475
Db 2281 AACATCATGAAACCCCGTCTCTACTTAAATAACAAAAATTAGCCAGGATGTTGGCACAC 2340
Qy 2476 ACCTGTAGTCCGACTACTCGGGAGCCCAACGACGAGAACCGCTTGTACCCAGGAGGTG 2535
Db 2341 ACCTGTAGTCCGACTACTCGGGAGCCCAACGACGAGAACCGCTTGTACCCAGGAGGTG 2400
Qy 2536 GAGGTTCGAGTGCAGCCGAAAGTTGCACATTGCACTTCCACCTCGGCGCAGACAGCAAGATT 2595
Db 2401 GAGGTTCGAGTGCAGCCGAAAGTTGCACATTGCACTTCCACCTCGGCGCAGACAGCAAGATT 2460
Qy 2596 CTATCAAAAAAAGGACGAGTGGCAAGTAAAGTATAGAGAGAAATGCTGTAGAGGAA 2655
Db 2461 CTATCAAAAAAAGGACGAGTGGCAAGTAAAGTATAGAGAGAAATGCTGTAGAGGAA 2520
Qy 2656 TTAAGCGTTGTAGTAAACGCGTCTCATCTCTAAGCTTTGAAGAGGAGAGCAAAAAATCC 2715
Db 2521 TTAAGCGTTGTAGTAAACGCGTCTCATCTCTAAGCTTTGAAGAGGAGAGCAAAAAATCC 2580
Qy 2716 ATTTGTTTAAATTCACATCTCAAGAGGAGGAAACCCGCGCTGTGTGGGTGTTGCCAAT 2775
Db 2581 ATTTGTTTAAATTCACATCTCAAGAGGAGGAAACCCGCGCTGTGTGGGTGTTGCCAAT 2640
Qy 2776 TTCCTAGAACGGAATGTGGGGTATAGAAAAAGGAATGAATTAAGCTGTTTTCAAAT 2835
Db 2641 TTCCTAGAACGGAATGTGGGGTATAGAAAAAGGAATGAATTAAGCTGTTTTCAAAT 2700
Qy 2836 AGGTTCCTTTGTAAGTTATTGATGAGAGGAAAGATTCAGCTGGGAGGCGCTTAAAAATGAT 2895
Db 2701 AGGTTCCTTTGTAAGTTATTGATGAGAGGAAAGATTCAGCTGGGAGGCGCTTAAAAATGAT 2760
Qy 2896 TTGGGAAACAAATTCCTTTGAGGCTCAGTGACAAACGCAAGATTAACAATT 2948
Db 2761 TTGGGAAACAAATTCCTTTGAGGCTCAGTGACAAACGCAAGATTAACAATT 2813

RESULT 3

AR371889 LOCUS 2813 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 3 from patent US 6395544.
ACCESSION AR371889
VERSION AR371889.1 GI:34608999
KEYWORDS SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2813)
AUTHORS Cowser, L.M. and Freier, S.M.
TITLE Antisense modulation of BCAS1 expression

JOURNAL Patent: US 639544-A 3 28-MAY-2002;

FEATURES

Location/Qualifiers

source

1..2813

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 91.7%; Score 2813; DB 6; Length 2813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 CTCACCCAGCAGCAGGCACTGGGCGACGACCTGGAGACCCAGGACCCCTGTGCGAGG 195
DB 1 CTCACCCAGCAGCAGGCACTGGGCGACGACCTGGAGACCCAGGACCCCTGTGCGAGG 60

QY 196 CAGTCCCGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGTCTCAGCAGGAGCAATG 255
DB 61 CAGTCCCGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGTCTCAGCAGGAGCAATG 120

QY 256 GGTACCAAAATGAGTGTCCCAAGAGTTGAAGACCAAGAGATGAACCAAGCAGACAG 315
DB 121 GGTACCAAAATGAGTGTCCCAAGAGTTGAAGACCAAGAGATGAACCAAGCAGACAG 180

QY 316 ACTTACCAAGCAACCGCTCTCTCTGAACGGGGTTCCAGTGGTGTCTGACACCCACACA 375
DB 181 ACTTACCAAGCAACCGCTCTCTCTGAACGGGGTTCCAGTGGTGTCTGACACCCACACA 240

QY 376 GTTCAGCACTTAGAGGAAGTGCATTTGGGAATAAGTGTCAAGACGGATTAATGTGCCACT 435
DB 241 GTTCAGCACTTAGAGGAAGTGCATTTGGGAATAAGTGTCAAGACGGATTAATGTGCCACT 300

QY 436 TCTTCCCGGAGACACCGAGATAGTCTGTTCGGATGTCACCGGAAGAGATCTTGGG 495
DB 301 TCTTCCCGGAGACACCGAGATAGTCTGTTCGGATGTCACCGGAAGAGATCTTGGG 360

QY 496 AAAGAGGCCAAACCCGAGCACCAGTCTGTAATCTCGTTTTTCTTGATGCTCTCTCGG 555
DB 361 AAAGAGGCCAAACCCGAGCACCAGTCTGTAATCTCGTTTTTCTTGATGCTCTCTCGG 420

QY 556 CTTGTACAGGACGTACCCGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTT 615
DB 421 CTTGTACAGGACGTACCCGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTT 480

QY 616 GATGTACGCTCCATTAAGCTCCAGCGAACCAAGACCCCAAGTGAGAGCTTGACATTCG 675
DB 481 GATGTACGCTCCATTAAGCTCCAGCGAACCAAGACCCCAAGTGAGAGCTTGACATTCG 540

QY 676 GTGGCAGCTGGACCGGGGAGGACACAGATAAACCCAGGAGCAGCCCGGCCCAAGAC 735
DB 541 GTGGCAGCTGGACCGGGGAGGACACAGATAAACCCAGGAGCAGCCCGGCCCAAGAC 600

QY 736 AAGTGTCTCTCTGCGCCAGGGATCCACGCTTCTCCACCTGAGACAGGGGGAGCAGGA 795
DB 601 AAGTGTCTCTCTGCGCCAGGGATCCACGCTTCTCCACCTGAGACAGGGGGAGCAGGA 660

QY 796 GGAGAGCTCCCTCCAGCCCAAGAGCTCCAGCTTTTGTGACAAATCTTCAAGCTGGAC 855
DB 661 GGAGAGCTCCCTCCAGCCCAAGAGCTCCAGCTTTTGTGACAAATCTTCAAGCTGGAC 720

QY 856 AAGGACAGGAAAGGTGCGAGTGCACAGCCAAACAGGAAGCCAGAGGGGAGAGCATCAA 915
DB 721 AAGGACAGGAAAGGTGCGAGTGCACAGCCAAACAGGAAGCCAGAGGGGAGAGCATCAA 780

QY 916 GACAAAGTGGATGAGGTTCTTGGCTTATCAGGGCAGTCCGATGATGTCCTGCGAGGAAG 975
DB 781 GACAAAGTGGATGAGGTTCTTGGCTTATCAGGGCAGTCCGATGATGTCCTGCGAGGAAG 840

QY 976 GACATAGTTGACGGCAAGAAAGAGCAAGAACTTTGGAACTTGCGGATGTCCTGTC 1035
DB 841 GACATAGTTGACGGCAAGAAAGAGCAAGAACTTTGGAACTTGCGGATGTCCTGTC 900

QY 1036 CTTGGGACCCAGAGGACTGGAGACTGCNAGGACGATTCACGAGCAGCAGCTATAGCA 1095

DB 901 CTTGGGACCCAGAGGACTGGAGACTGCNAGGACGATTCCTCCAGGACGAGCTATAGCA 960

QY 1096 GAGATAATAATTCATCATGAGTTCTTTTAAACTCTGGTTTACCTTAACAAAGCTGAA 1155

DB 961 GAGATAATAATTCATCATGAGTTCTTTTAAACTCTGGTTTACCTTAACAAAGCTGAA 1020

QY 1156 ACAAAAAGGACCCAGNAGACACGGGTCTGAAAAGTCAACCCACCTTCAGCTGACCTT 1215

DB 1021 ACAAAAAGGACCCAGNAGACACGGGTCTGAAAAGTCAACCCACCTTCAGCTGACCTT 1080

QY 1216 AAGTCAGACAAAGCCAACTTTTACATCCAGGAGACCCAAAGGGGCTGGCAAGAAATTCCAA 1275

DB 1081 AAGTCAGACAAAGCCAACTTTTACATCCAGGAGACCCAAAGGGGCTGGCAAGAAATTCCAA 1140

QY 1276 GGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTGAACTTCGCGAAGAGGC 1335

DB 1141 GGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTGAACTTCGCGAAGAGGC 1200

QY 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAGTCA 1395

DB 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAGTCA 1260

QY 1396 GTTAAAGAGGACTCAGTCCCCACAGGTGGGAGAGAAATGTGGTGTGTGAGTCACCACTA 1455

DB 1261 GTTAAAGAGGACTCAGTCCCCACAGGTGGGAGAGAAATGTGGTGTGTGAGTCACCACTA 1320

QY 1456 GAGATTATAAGTCCAGGAAAGTGAATCAGCTTTTCAAAAGTGGAACTTCAACGAAGA 1515

DB 1321 GAGATTATAAGTCCAGGAAAGTGAATCAGCTTTTCAAAAGTGGAACTTCAACGAAGA 1380

QY 1516 GATCTGCACCTGACACCCACAGGACGAAACTCAAAAGAGAGAAAGCAACCAAGAAC 1575

DB 1381 GATCTGCACCTGACACCCACAGGACGAAACTCAAAAGAGAGAAAGCAACCAAGAAC 1440

QY 1576 TCTCTGATGGGTTTCTCAGACAAATGTCAAGTGAAGGGGATGGAGGATCAACCACTCA 1635

DB 1441 TCTCTGATGGGTTTCTCAGACAAATGTCAAGTGAAGGGGATGGAGGATCAACCACTCA 1500

QY 1636 GAAGAAATAATGGGAAAGACTCAGCTGCGCAACATCAGACTCCACAGAAAAGACTATC 1695

DB 1501 GAAGAAATAATGGGAAAGACTCAGCTGCGCAACATCAGACTCCACAGAAAAGACTATC 1560

QY 1696 ACACCGCAGAGCTGACCAACAGGAGCACCACAGAGGGTAAAGAGGCTCTCGAG 1755

DB 1561 ACACCGCAGAGCTGACCAACAGGAGCACCACAGAGGGTAAAGAGGCTCTCGAG 1620

QY 1756 GACAAGAGTCAAGCAGCCGAGATGAACAAAGCAGAGAGCAACCAAGCAGAGGACCAAGAA 1815

DB 1621 GACAAGAGTCAAGCAGCCGAGATGAACAAAGCAGAGAGCAACCAAGCAGAGGACCAAGAA 1680

QY 1816 CCAGCCAGTGCACAGCAGGCCACGTTGGACACGAACTCACTCAGAAATGGGACAAAG 1875

DB 1681 CCAGCCAGTGCACAGCAGGCCACGTTGGACACGAACTCACTCAGAAATGGGACAAAG 1740

QY 1876 CTCNAAAGAGACCTGAGAGGGGCGAGTCCCTTGGGGGCTCTTTAAAGGCTCGGA 1935

DB 1741 CTCNAAAGAGACCTGAGAGGGGCGAGTCCCTTGGGGGCTCTTTAAAGGCTCGGA 1800

QY 1936 CCAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGATATCCATCGGACAGTTGGC 1995

DB 1801 CCAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGATATCCATCGGACAGTTGGC 1860

QY 1996 AAACCCAGTAAACAAATCAGACCGGTTCCCAACAGGTTCTCTGCCACCAAGATGTGTT 2055

DB 1861 AAACCCAGTAAACAAATCAGACCGGTTCCCAACAGGTTCTCTGCCACCAAGATGTGTT 1920

QY 2056 CTCCTTACTTCATCTCTCCCAACACGCTCCATGATATATTTCTTCTGATGGCAGCA 2115

DB 1921 CTCCTTACTTCATCTCTCCCAACACGCTCCATGATATATTTCTTCTGATGGCAGCA 1980

QY 2116 AATGAAATTCCTGCTAGAAATTAAGCCCGAGCTGTTGATATTTAGGTTGATTTTACG 2175

DB 1981 AATGAAATTCCTGCTAGAAATTAAGCCCGAGCTGTTGATATTTAGGTTGATTTTACG 2040

```
Qy 2176 TCTCTGGTCCAGTCTTTTCTGGCAATAACAGTAAAGATGTTTGTAGCAGGTCACTAGTT 2235
Db 2041 TCTCTGGTCCAGTCTTTTCTGGCAATAACAGTAAAGATGTTTGTAGCAGGTCACTAGTT 2100
Qy 2236 GGGTCAGAAAGTCATGATCACCAGCAGAAAGGGAGGAATAGAGGAATGTGTCGG 2295
Db 2101 GGGTCAGAAAGTCATGATCACCAGCAGAAAGGGAGGAATAGAGGAATGTGTCGG 2160
Qy 2296 GTTAAGTCATAAATGGCAGTGTGGCGGGCTGTGGCTCTCGCTGTAATCTCAGC 2355
Db 2161 GTTAAGTCATAAATGGCAGTGTGGCGGGCTGTGGCTCTCGCTGTAATCTCAGC 2220
Qy 2356 ACTTTGGAGGCCAGGAGGTGATCACCCTGAGTCAGAGGTTCAGAGCTAGCTGGCC 2415
Db 2221 ACTTTGGAGGCCAGGAGGTGATCACCCTGAGTCAGAGGTTCAGAGCTAGCTGGCC 2280
Qy 2416 AACATCATGAACCCCGTCTCTACTAAATAATACAAAATATGACGAGCATCGTGGCACAC 2475
Db 2281 AACATCATGAACCCCGTCTCTACTAAATAATACAAAATATGACGAGCATCGTGGCACAC 2340
Qy 2476 ACCTGTAGTCCAGTACTCGGAGGCCCAACGCACGAGAACCGCTTGTACCCAGGAGTG 2535
Db 2341 ACCTGTAGTCCAGTACTCGGAGGCCCAACGCACGAGAACCGCTTGTACCCAGGAGTG 2400
Qy 2536 GAGGTTCAGTCAGCCGAAGTTGCACCATTTGCACCTCCACCTCGGCGCACAGCAAGATT 2595
Db 2401 GAGGTTCAGTCAGCCGAAGTTGCACCATTTGCACCTCCACCTCGGCGCACAGCAAGATT 2460
Qy 2596 CTATCAAAAAAAGGAGGAGTGAAGTTATAGAAAGAAATGCTGTAGAGGAA 2655
Db 2461 CTATCAAAAAAAGGAGGAGTGAAGTTATAGAAAGAAATGCTGTAGAGGAA 2520
Qy 2656 TTAAGCGTTGAGTAAACCGCTGCTCATCTTAAGCTTGAAGAGGAGGAGCAAAATCC 2715
Db 2521 TTAAGCGTTGAGTAAACCGCTGCTCATCTTAAGCTTGAAGAGGAGGAGCAAAATCC 2580
Qy 2716 ATTTGTTTAAATTCACATCTCAAGAGGAGGAGAACCCCGGCTGTGTGGTGGTGGCCAAAT 2775
Db 2581 ATTTGTTTAAATTCACATCTCAAGAGGAGGAGAACCCCGGCTGTGTGGTGGTGGCCAAAT 2640
Qy 2776 TTCTTAGAACCGAATGTGTGGGTATAGAAAAGGAATGAATAAGCGTTGTTTCAAT 2835
Db 2641 TTCTTAGAACCGAATGTGTGGGTATAGAAAAGGAATGAATAAGCGTTGTTTCAAT 2700
Qy 2836 AGGTCTCTGTAAGTTATGATGAGGAGGAGAAAGATTGACTGGGAGGCTTAAATCAT 2895
Db 2701 AGGTCTCTGTAAGTTATGATGAGGAGGAGAAAGATTGACTGGGAGGCTTAAATCAT 2760
Qy 2896 TTGGGAAAAAATTCCTTTTGAAGGCTCAGTGACACCGCAAGATTGACAACTT 2948
Db 2761 TTGGGAAAAAATTCCTTTTGAAGGCTCAGTGACACCGCAAGATTGACAACTT 2813

RESULT 4
LOCUS AF041260 2813 bp mRNA linear PRI 22-JUL-1998
DEFINITION Homo sapiens AIB1 (AIB1) mRNA, complete cds.
ACCESSION AF041260
VERSION AF041260.1 GI:3335398
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2813)
AUTHORS Collins,C., Rommens,J.M., Kowbel,D., Godfrey,T., Tanner,M.,
Hwang,S.-I., Polikoff,D., Nonet,G., Cochran,J., Myambo,K.,
Jay,K.E., Froula,J., Cloutier,T., Kuo,W.-L., Yaswen,P., Dairkee,S.,
Giovannola,J., Hutchinson,G.B., Isola,J., Kallioniemi,O.-P.,
Palazzolo,M., Martin,C., Ericsson,C., Pinkel,D., Albertson,D.,
Li,W.-B. and Gray,J.W.
Positional cloning of ZNF217 and NAB1: genes amplified at 20q13.2
```

```
JOURNAL and overexpressed in breast carcinoma
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 95 (15), 8703-8708 (1998)
PUBMED 98337979
REFERENCE 9671742
AUTHORS 2 (bases 1 to 2813)
Collins,C., Rommens,J.M., Kowbel,D., Godfrey,T., Tanner,M.,
Hwang,S.-I., Polikoff,D., Nonet,G., Cochran,J., Myambo,K.,
Jay,K.E., Froula,J., Cloutier,T., Kuo,W.-L., Yaswen,P., Dairkee,S.,
Giovannola,J., Hutchinson,G.B., Isola,J., Kallioniemi,O.-P.,
Palazzolo,M., Martin,C., Ericsson,C., Pinkel,D., Albertson,D.,
Li,W.-B. and Gray,J.W.
Direct Submission
Submitted (07-JAN-1998) Life Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
FEATURES
source
1..2813
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.2-q13.3"
1..2813
/gene="AIB1"
118..1872
/gene="AIB1"
/ncbi_start=1
/ncbi_end=1872
/product="AIB1"
/protein_id="AAC39896.1"
/db_xref="GI:3335398"
/translation="MGNQMSVPQVDEONEPEABTYODNASALNGVPPVVVSTHTVQH
LEEVLDGISVKTNDVNTSSPETTEISAVADANGKLGKEAPEAKSRFLMLSRP
VPGRTGQAADSSLGSLVSSNKAPKDPSSWTLPAVAGPQDQTKTPGHAPAQ
DKVLSAARDPTLLPPTGGAGGAPSKDSSPFDFKLDKGQKVPDQSQAKRA
EHQKXDEVPLSGSDDDVPAGKDIVDGEKEGQELCTADCVSDPDEGLTAKDDSO
AAATANNNSIMSFFKTLVSPKAEKTKDPEDEAKSPPTSALDKSKDNFTSQETQ
GAGNKGKNSPMSFTVTPPEAKETKPEKSGPTSLPLGLFWKKSXKEDSVPTGAE
ENVVCEGPVRIIKSEVESALQTVLDNEGDAAPTEAKLREESKPTSLSMAFLRQM
SUKGDIITHEEINGKSSQTSQSDTEKTTTPPEPTGAPQKKGSSDKKSAAE
MNKQSNKORAKEPAQCTEQATVDNTSLQNGDKLQKRPKRSQSLGGFFKGLGPKRML
DAQVQTPVSLGVPVKPK"
```

ORIGIN

```
Query Match 91.7%; Score 2813; DB 9; Length 2813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 136 CTCACCCGACAGCAGGACCTGGGCGAGCACCGACCTGGAGACCCAGGACCTGTGCAGGAG 195
Db 1 CTCACCCGACAGCAGGACCTGGGCGAGCACCGACCTGGAGACCCAGGACCTGTGCAGGAG 60
Qy 196 CAGTCCGGGTGACACGAGGGGACTGGAAGTACTCCACAGGGGCTCAGCAGGAGCAATG 255
Db 61 CAGTCCGGGTGACACGAGGGGACTGGAAGTACTCCACAGGGGCTCAGCAGGAGCAATG 120
Qy 256 GGTAACCAATAGAGTGTTCCTCCCAAGAGTGTGAAGCAAGAGATGAACCAAGAGCAGAG 315
Db 121 GGTAACCAATAGAGTGTTCCTCCCAAGAGTGTGAAGCAAGAGATGAACCAAGAGCAGAG 180
Qy 316 ACTTACAGGACACCGGCTGTCTGTAACGGGGTTCAGTGGTGGTGTGACCCACACA 375
Db 181 ACTTACAGGACACCGGCTGTCTGTAACGGGGTTCAGTGGTGGTGTGACCCACACA 240
Qy 376 GTTCAGCACTTAGAGGAAGTCAGCTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 435
Db 241 GTTCAGCACTTAGAGGAAGTCAGCTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 300
Qy 436 TCTTCCCCCGAGACCAACCGAGATAAGTCTGTTCGGATGCCAACCGAAAGAAATCTTGGG 495
Db 301 TCTTCCCCCGAGACCAACCGAGATAAGTCTGTTCGGATGCCAACCGAAAGAAATCTTGGG 360
Qy 496 AAAGAGGCCAAACCCGAGGACCGAGCTCTAAATCTCGTTTTTTTCTTCTGCTCTCGG 555
```

Db 361 AAGAGGCCAACCAGGACCCAGCTGCTAAATCTCGTTTTTTTCTTGATGCTCTCTCGG 420
Qy 556 CCGTGTACAGAGAGCTACCGGAGACCAAGCCGAGATTCAATCCCTTGGATTCAGTGAAGCTT 615
Db 421 CCGTGTACAGAGAGCTACCGGAGACCAAGCCGAGATTCAATCCCTTGGATTCAGTGAAGCTT 480
Qy 616 GATGTCAAGTCCCAATTAAGTCTCCAGGACCAAGAACCCCAAGTGAGAGCTGGACACTCCG 675
Db 481 GATGTCAAGTCCCAATTAAGTCTCCAGGACCAAGAACCCCAAGTGAGAGCTGGACACTCCG 540
Qy 676 GTGGCAGCTGGACCGGGGAGGACACAGATAAACCACAGGGGACCGCCCGGCCCAAGAC 735
Db 541 GTGGCAGCTGGACCGGGGAGGACACAGATAAACCACAGGGGACCGCCCGGCCCAAGAC 600
Qy 736 AAGTCTCTCTGTCCGCGCAGGATCCAGCTTCTCCACTGTGAGACAGGGGAGGACGGA 795
Db 601 AAGTCTCTCTGTCCGCGCAGGATCCAGCTTCTCCACTGTGAGACAGGGGAGGACGGA 660
Qy 796 GGAGAAGCTCCCTCCAAAGCCCAAGACTCCAGCTTTTGTGACAAATCTTCAAGCTGGAC 855
Db 661 GGAGAAGCTCCCTCCAAAGCCCAAGACTCCAGCTTTTGTGACAAATCTTCAAGCTGGAC 720
Qy 856 AAGGACACAGGAAAAGGTCCAGTGACAGCCAAAGGAAAGCCCAAGAGGCGCAGAGCATCAA 915
Db 721 AAGGACACAGGAAAAGGTCCAGTGACAGCCAAAGGAAAGCCCAAGAGGCGCAGAGCATCAA 780
Qy 916 GACAGGTGGATGAGGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCAAGGGAAG 975
Db 781 GACAGGTGGATGAGGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCAAGGGAAG 840
Qy 976 GACATAGTTGACGCAAGGAAAAGGACAGGACAGAACTTGGAACTGGCGGATGCTCTGTC 1035
Db 841 GACATAGTTGACGCAAGGAAAAGGACAGGAACTTGGAACTGGCGGATGCTCTGTC 900
Qy 1036 CCGTGGGACCCAGAGGACTGGAGACTGCAAGAGACGATTCGCCAGGACGAGCTATAGCA 1095
Db 901 CCGTGGGACCCAGAGGACTGGAGACTGCAAGAGACGATTCGCCAGGACGAGCTATAGCA 960
Qy 1096 GAGATAATAATTCATCATGATGTTCTTTAAACTCTGGTTTCACTTAAACAAAGCTGAA 1155
Db 961 GAGATAATAATTCATCATGATGTTCTTTAAACTCTGGTTTCACTTAAACAAAGCTGAA 1020
Qy 1156 ACNAAAAGGACCCAGAGACACGGTGTGAAAAGTCAACCACTTCACTGACCTTCACTGACCTT 1215
Db 1021 ACNAAAAGGACCCAGAGACACGGTGTGAAAAGTCAACCACTTCACTGACCTTCACTGACCTT 1080
Qy 1216 AAGTCAGACAAAGCCAACTTTACATCCAGAGACCCCAAGGGCTGGCAAGATTTCCAAA 1275
Db 1081 AAGTCAGACAAAGCCAACTTTACATCCAGAGACCCCAAGGGCTGGCAAGATTTCCAAA 1140
Qy 1276 GGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTGAACTTCGGAAGGAAGGC 1335
Db 1141 GGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTGAACTTCGGAAGGAAGGC 1200
Qy 1336 ACCAAGGAGAAATCAGGACCACTCTCTGCTCTGGGCAAACTGTTTGGAAAAGTCA 1395
Db 1201 ACCAAGGAGAAATCAGGACCACTCTCTGCTCTGGGCAAACTGTTTGGAAAAGTCA 1260
Qy 1396 GTTAAAGAGGACTCAGTCCCAAGTCCGAGGAGAAATGTTGTTGAGTCAACAGTA 1455
Db 1261 GTTAAAGAGGACTCAGTCCCAAGTCCGAGGAGAAATGTTGTTGAGTCAACAGTA 1320
Qy 1456 GAGATTATAAGTCCCAAGGAGTAGAATCAGCTTTTACAAACAGTGGACCTCAACGAAGGA 1515
Db 1321 GAGATTATAAGTCCCAAGGAGTAGAATCAGCTTTTACAAACAGTGGACCTCAACGAAGGA 1380
Qy 1516 GATCTGCACTGAAACCCACAGAGCGGAAACTCAAAAGAGAGAAAGCAAAACCAAGAAC 1575
Db 1381 GATCTGCACTGAAACCCACAGAGCGGAAACTCAAAAGAGAGAAAGCAAAACCAAGAAC 1440
Qy 1576 TCTCTGATGGCTTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGATCAACCCACTCA 1635
Db 1441 TCTCTGATGGCTTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGATCAACCCACTCA 1500

Qy 1636 GAAGAAATAATGGGAAAGACTCCAGTGTCCAAACATCAGACTCCACAGAAAAGACTATC 1695
Db 1501 GAAGAAATAATGGGAAAGACTCCAGTGTCCAAACATCAGACTCCACAGAAAAGACTATC 1560
Qy 1696 ACACGCCAGAGCTGTGAACCAACAGGAGCACCAAGAGGGTAAAGAGGGCTCCTCGAAG 1755
Db 1561 ACACGCCAGAGCTGTGAACCAACAGGAGCACCAAGAGGGTAAAGAGGGCTCCTCGAAG 1620
Qy 1756 GACAAGAGTCCAGCAGCGGAGATGAACAGCAGAGAGCAACAGAGGAGGACCAAGAA 1815
Db 1621 GACAAGAGTCCAGCAGCGGAGATGAACAGCAGAGAGCAACAGAGGAGGACCAAGAA 1680
Qy 1816 CCAGCCCATGTCCACAGAGCAGGCGCAGGTGACACGAACTCACTGCAAGATGGGGACAAG 1875
Db 1681 CCAGCCCATGTCCACAGAGCAGGCGCAGGTGACACGAACTCACTGCAAGATGGGGACAAG 1740
Qy 1876 CTCCAAAGAGACCTGTGAAGCGGACAGTCCCTTGGGGCTCTTTTAAAGGCTCGGA 1935
Db 1741 CTCCAAAGAGACCTGTGAAGCGGACAGTCCCTTGGGGCTCTTTTAAAGGCTCGGA 1800
Qy 1936 CCAAAGCGGATGTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGACCAAGTGGC 1995
Db 1801 CCAAAGCGGATGTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGACCAAGTGGC 1860
Qy 1996 AAAACCAAGTAAACAAATCAGCACGGTTCCACAGGTTCTCTGCCACCAAGATGTGT 2055
Db 1861 AAAACCAAGTAAACAAATCAGCACGGTTCCACAGGTTCTCTGCCACCAAGATGTGT 1920
Qy 2056 CTCCTTACTCCATCTCTCCCAACAGCTCCATGTATATATTTCTTGTGATGGCCAGA 2115
Db 1921 CTCCTTACTCCATCTCTCCCAACAGCTCCATGTATATATTTCTTGTGATGGCCAGA 1980
Qy 2116 AATGAAATTCCTGCTAGAAATTAAGCCGAGCTGTGTATATTGAGGTGATTTATTACG 2175
Db 1981 AATGAAATTCCTGCTAGAAATTAAGCCGAGCTGTGTATATTGAGGTGATTTATTACG 2040
Qy 2176 TCTCTGTCTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTGTAGCAGGTCACTAGT 2235
Db 2041 TCTCTGTCTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTGTAGCAGGTCACTAGT 2100
Qy 2236 GGGTCAGAGAGTGCATGATCACCAGCAGGAAAGGAGGGAATAGAGGAATGTTTCGG 2295
Db 2101 GGGTCAGAGAGTGCATGATCACCAGCAGGAAAGGAGGGAATAGAGGAATGTTTCGG 2160
Qy 2296 GTTAAAGTGAATAATGGCAGTGTGCGCGGCGTGTGGCTCTCGCTGTAATCTCAGC 2355
Db 2161 GTTAAAGTGAATAATGGCAGTGTGCGCGGCGTGTGGCTCTCGCTGTAATCTCAGC 2220
Qy 2356 ACTTTGGAGGCCGAGCAGGTGGATCACTTGAGTCAAGAGTTCAAGACTAGCTTGCC 2415
Db 2221 ACTTTGGAGGCCGAGCAGGTGGATCACTTGAGTCAAGAGTTCAAGACTAGCTTGCC 2280
Qy 2416 AACATCATGAACCCCGTCTCTACTATAAATAACAAAAATAGCCAGGATGGTGGCACAC 2475
Db 2281 AACATCATGAACCCCGTCTCTACTATAAATAACAAAAATAGCCAGGATGGTGGCACAC 2340
Qy 2476 ACCTGTAGTCCAGCTACTCGGGAGCCCAACGACAGAGACCGCTTGTACCCAGAGGTG 2535
Db 2341 ACCTGTAGTCCAGCTACTCGGGAGCCCAACGACAGAGACCGCTTGTACCCAGAGGTG 2400
Qy 2536 GAGTTGAGTGGACCCGAAGTTGCACCTTGTCACTCCACCTCGGGCGCAGAGCAAGATT 2595
Db 2401 GAGTTGAGTGGACCCGAAGTTGCACCTTGTCACTCCACCTCGGGCGCAGAGCAAGATT 2460
Qy 2596 CTATCAAAAAAAGGAGCAGTGGCAAGTATAGAGAGAAATGCTGTAGAGGAA 2655
Db 2461 CTATCAAAAAAAGGAGCAGTGGCAAGTATAGAGAGAAATGCTGTAGAGGAA 2520
Qy 2656 TTAAGCGTTGTAGTAAACCGGTGCTCATCTCTTAAGCTTGAAGAGGAGACGAAAAATCC 2715
Db 2521 TTAAGCGTTGTAGTAAACCGGTGCTCATCTCTTAAGCTTGAAGAGGAGACGAAAAATCC 2580

1636 GAAGAAATAAATGGGAAGACTCCAGCTGCCAAAACATCAGACTCCACAGAAAAGACTATC 1695
1493 GAAGAAATAAATGGGAAGACTCCAGCTGCCAAAACATCAGACTCCACAGAAAAGACTATC 1552
1696 ACACCGCCAGAGCCTGAAACCAACAGAGAGCACCACAGAGGGGTAAAGAGGGCTTCCTCGAAG 1755
1553 ACACCGCCAGAGCCTGAAACCAACAGAGAGCACCACAGAGGGGTAAAGAGGGCTTCCTCGAAG 1612
1756 GACAGAAGTACAGAGCCGAGATGAACAAAGAGAGAGCAACAGAGGAAGCCAAAGAA 1815
1613 GACAGAAGTACAGAGCCGAGATGAACAAAGAGAGAGCAACAGAGGAAGCCAAAGAA 1672
1816 CCAGCCCAAGTGCACAGAGCAGCCAGCGTGGACACGAACTCACTCGAATGGGACCAAG 1875
1673 CAGGCCCAAGTGCACAGAGCAGCCAGCGTGGACACGAACTCACTCGAATGGGACCAAG 1732
1876 CTCCTAAAGAGAGCCTGAGAAGCGGAGCAGAGTCCCTTTGGGGGCTCTTTAAAGCCCTGGGA 1935
1733 CTCCTAAAGAGAGCCTGAGAAGCGGAGCAGAGTCCCTTTGGGGGCTCTTTAAAGCCCTGGGA 1792
1936 CCAAGCGGATGTTGATGCTCAAGTGCAACAGAGCCAGTATCATCGGACCAAGTTGGC 1995
1793 CCAAGCGGATGTTGATGCTCAAGTGCAACAGAGCCAGTATCATCGGACCAAGTTGGC 1852
1996 AAACCCCAAGTAAACAAATCAGCACGCTTCCACACAGGTTCTCCTGCCACCAAGATGTCT 2055
1853 AAATCAAGTAAACAAATCAGCACGCTTCCACACAGGTTCTCCTGCCACCAAGATGTCT 1912
2056 CTCCTTACTCTCCCTCCCAACACAGCTCCATGTATATATCTCTGATGGCCAGCA 2115
1913 CTCCTTACTCTCCCTCCCAACACAGCTCCATGTATATATCTCTGATGGCCAGCA 1972
2116 AATGAAATTCGCTAGAAATTAAGCCGAGCTGTGTATATATGAGGTGTATATTACG 2175
1973 AATGAAATTCGCTAGAAATTAAGCCGAGCTGTGTATATATGAGGTGTATATTACG 2032
2176 TCTCTGTCCTCAGTCTTTCTGCAATTAACAGTAAAGATGTTTACGAGTCACTAGT 2235
2033 TCTCTGTCCTCAGTCTTTCTGCAATTAACAGTAAAGATGTTTACGAGTCACTAGT 2092
2236 GGGTCAGAGAGTCGATGATACCAAGCAGAGGAAGGGAGGAATAGAGGAATGTCTCGG 2295
2093 GGGTCAGAGAGTCGATGATACCAAGCAGAGGAAGGGAGGAATAGAGGAATGTCTCGG 2152
2296 GTTAAGTATGAAATGAGCAGTGTGTGGCGGCGTGTGGCTCTCGCTGTAAATCTCAGC 2355
2153 GTTAAGTATGAAATGAGCAGTGTGTGGCGGCGTGTGGCTCTCGCTGTAAATCTCAGC 2212
2356 ACTTTGGGAGCGGAGGAGGAGTGGATCACCTGAGGTCAGGATTCAGAGTCAAGCTAGCCTGCC 2415
2213 ACTTTGGGAGCGGAGGAGGAGTGGATCACCTGAGGTCAGGATTCAGAGTCAAGCTAGCCTGCC 2272
2416 AACATCATGAACCCCGTCTCTACTAAATAACAAAATTAAGCAGGAGTGTGGCCACAC 2475
2273 AACATCATGAACCCCGTCTCTACTAAATAACAAAATTAAGCAGGAGTGTGGCCACAC 2332
2476 ACTCTAGTCCAGTACTCTGGAGGCCCAACGCAAGAGCCGCTTTGATCCAGGAGGTG 2535
2333 ACTCTAGTCCAGTACTCTGGAGGCCCAACGCAAGAGCCGCTTTGATCCAGGAGGTG 2392
2536 GAGGTTGAGTGAAGGAGTGGACCATTCAGCTCCACCTGGGCGGACAGGCAAGATT 2595
2393 GAGGTTGAGTGAAGGAGTGGACCATTCAGCTCCACCTGGGCGGACAGGCAAGATT 2452
2596 CTATC- - -AAAAAAGAGGAGTGGCAGTAAAGTATAGAGAGAAATGCTGTAGAG 2652
2453 CTATCAAAAAAAGAGGAGTGGCAGTAAAGTATAGAGAGAAATGCTGTAGAG 2512
2653 GAATTAAGCGTTGTAGTAAACCGCTGCTCATCTCTTAAGCTTTGAAGAGGGAGCAAAA 2712
2513 GAATTAAGCGTTGTAGTAAACCGCTGCTCATCTCTTAAGCTTTGAAGAGGGAGCAAAA 2572
2713 TCCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTTGGTGGTTGCC 2772

Db 2573 TCCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTTGGTGGTTGCC 2632
Qy 2773 AATTTCTTAGAACCGAATGTGTGGGTATAGAAAAAGGAATGAATPAAGCGTTGTTTTTCA 2832
Db 2633 AATTTCTTAGAACCGAATGTGTGGGTATAGAAAAAGGAATGAATPAAGCGTTGTTTTTCA 2692
Qy 2833 AATAGGTCCTTGTAAAGTATTATGATGAGAGGAAAAAGATTGACTGGGAGGCTTAAAAAT 2892
Db 2693 AATAGGTCCTTGTAAAGTATTATGATGAGAGGAAAAAGATTGACTGGGAGGCTTAAAAAT 2752
Qy 2893 GATTTGGGAAAAAATTCCTTTTGGAGCTCAGTGACACGCAACAGATTACAACTT 2948
Db 2753 GATTTGGGAAAAAATTCCTTTTGGAGCTCAGTGACACGCAACAGATTACAACTT 2808

RESULT 6
AC005220
LOCUS Homo sapiens chromosome 20, BAC clone 99 (LBNL H80), complete
DEFINITION sequence.
ACCESSION AC005220
VERSION AC005220.1 GI:3282159
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 163196)
AUTHORS Gray, J.W., Collins, C., Kimmerly, W., Bondoc, M., Cheng, J., Connolly, K.S., Gunning, K.M., Kadner, K., Miguel, T., Miller, C., Piclucck, S., Pollard, M., Rojeski, H., Subramanian, S. and Martin, C.H.
TITLE Sequencing of human chromosome 20
REFERENCE 2 (bases 1 to 163196)
AUTHORS Ricke, D.O.
TITLE Large Scale Sequence Analysis and Annotation with the Sequence Comparison Analysis (SCAN) System
REFERENCE 3 (bases 1 to 163196)
AUTHORS Gray, J.W., Collins, C., Kimmerly, W., Bondoc, M., Cheng, J., Connolly, K.S., Gunning, K.M., Davis, C.A., Kadner, K., Miguel, T., Piclucck, S., Pollard, M., Rojeski, H., Subramanian, S. and Martin, C.H.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A.
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES
source
1. 163196
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20"
/clone="BAC 99"
/note="LBNL H80"
4588..4709
/rpt_family="MIR"
6685..6813
/rpt_family="Tigger1"
complement(6838..7220)
/rpt_family="L1"
7175..7314
/rpt_family="L1"
7322..7441
/rpt_family="Tigger1"
complement(8695..9257)
/rpt_family="MLT2B2"
10121..10350
/rpt_family="Alu"
11057..11357
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region

```
repeat_region /rpt_family="Alu"
13431. .13702
/rpt_family="Alu"
repeat_region 15911. .15962
/rpt_family="MIR"
misc_feature 16029. .16484
/note="99% identity B54864 BAC end sequence"
repeat_region /rpt_family="Alu"
16891. .17025
/rpt_family="Alu"
repeat_region /rpt_family="L1"
17221. .17682
/rpt_family="L1"
repeat_region /rpt_family="L1"
18500. .18756
/rpt_family="L1"
misc_feature /note="GRAIL 2 excellent exon, frame 1"
19640. .19943
/rpt_family="Alu"
repeat_region /rpt_family="Alu"
22025. .22314
/rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 1"
22903. .23020
/rpt_family="L1"
repeat_region /rpt_family="Alu"
23247. .23567
/rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 0"
24441. .24514
/rpt_family="L1"
repeat_region /note="GRAIL 2 excellent exon, frame 0"
25506. .25525
/rpt_type=tandem
/rpt_unit="t"
repeat_region /rpt_family="MIR"
25538. .25679
/rpt_family="MIR"
repeat_region /rpt_family="L1"
26920. .27577
/rpt_family="L1"
repeat_region /rpt_family="L1"
28412. .28563
/rpt_family="MIR"
28734. .29019
/rpt_family="Alu"
repeat_region /rpt_family="THE1"
30266. .30455
/rpt_family="THE1"
misc_feature 30789. .30888
/note="GRAIL 2 excellent exon, frame 2"
31841. .31972
/rpt_family="L1"
misc_feature /note="GRAIL 2 excellent exon, frame 0"
32270. .32290
/rpt_type=tandem
/rpt_unit="t"
32702. .32860
/note="GRAIL 2 excellent exon, frame 1"
33268. .33474
/rpt_family="L1"
repeat_region /note="GRAIL 2 excellent exon, frame 0"
33836. .33863
/rpt_type=tandem
/rpt_unit="t"
34144. .34436
/rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 0"
35158. .35265
/rpt_type=tandem
35605. .35855
/rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 1"
43253. .43465
/rpt_family="THE1"
repeat_region /rpt_family="L1"
50600. .50600
/rpt_family="L1"
repeat_region /rpt_family="L1"
50855. .50855
/rpt_family="L1"
repeat_region /rpt_family="THE1"
51461. .51822
/rpt_family="THE1"
repeat_region /rpt_family="L1"
54856. .54856
/rpt_family="L1"
repeat_region /rpt_family="L1"
55490. .55796
/rpt_family="Alu"
repeat_region 55900. .56027
/rpt_family="MIR"
misc_feature complement (61007. .61096)
/note="GRAIL 2 excellent exon, frame 1"
join (61724. .61846, 71254. .71319, 75055. .75261)
/standard_name="af74a01"
/note="100% identity dbEST:AA644590"
complement (62247. .62746)
/rpt_family="L1"
repeat_region /rpt_family="Alu"
65587. .65882
/rpt_family="Alu"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
68256. .68282
/rpt_type=tandem
/rpt_unit="t"
repeat_region /rpt_family="Alu"
69077. .69495
/rpt_family="Alu"
repeat_region /rpt_family="Alu"
69537. .69618
/rpt_family="Alu"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
70542. .70721
/rpt_family="MIR"
72072. .72255
/rpt_family="MIR"
repeat_region /rpt_family="MIR"
72260. .72340
/rpt_family="MIR"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
74417. .74456
/note="CT" 20"
/rpt_type=tandem
/rpt_unit="t"
75057. .75320
/note="GRAIL 2 excellent exon, frame 2"
join (75199. .75323, 84601. .84746)
/standard_name="zm96b07"
/note="96% identity dbEST:AA079436"
complement (77059. .77202)
/rpt_family="MIR"
repeat_region /rpt_family="MIR"
77269. .77551
/rpt_family="Alu"
prim_transcript join (78685. .79043, 88203. .88578)
/standard_name="ym20b03"
/note="97% identity dbEST:H15040"
79526. .79590
/standard_name="nc17d10"
/note="81% identity dbEST:AA259187"
complement (81547. .81627)
/rpt_family="MIR"
repeat_region 82700. .82865
/rpt_family="Alu"
repeat_region 83191. .83347
/rpt_family="L1"
repeat_region /note="GRAIL 2 excellent exon, frame 0"
83468. .83594
/rpt_family="MIR"
repeat_region 84144. .84436
/rpt_family="Alu"
prim_transcript complement (84402. .84774)
/standard_name="nl14b12"
/note="98% identity dbEST:AA554549"
complement (85065. .85162)
/standard_name="yw21h10"
/note="90% identity dbEST:H88501"
join (85165. .85589, 86249. .86579, 87354. .87476)
/standard_name="yg09d02"
/note="97% identity dbEST:R43020"
complement (86451. .86577)
/note="GRAIL 2 excellent exon, frame 2"
complement (87953. .88445)
/standard_name="zu90c12"
/note="99% identity dbEST:AA626177"
complement (89695. .89866)
/note="GRAIL 2 excellent exon, frame 1"
89959. .90146
/rpt_family="MIR"
```

```
Query Match      33.2%; Score 1016.4; DB 9; Length 163196;
Best Local Similarity 99.1%; Pred. No. 2.7e-270;
Matches 1033; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 1929 CCTGGGACCAACCGGATGTTGGATGCTCAAGTGAACACAGACCCAGTATCCATCGGACC 1988
DB 83752 CCAGGACCAACCGGATGTTGGATGCTCAAGTGAACACAGACCCAGTATCCATCGGACC 83811
QY 1989 AGTTGGCAACCAAGTAACAAATCAGCAGCGTTCCACACAGGTTCTCTGCGCAACGAG 2048
DB 83812 AGTTGGCAACCAAGTAACAAATCAGCAGCGTTCCACACAGGTTCTCTGCGCAACGAG 83871
QY 2049 ATGTGTTCTCTTACTTCCATCTCTCTCCCAACACGCTCCATGTATATATCTTCTGTATG 2108
DB 83872 ATGTGTTCTCTTACTTCCATCTCTCTCCCAACACGCTCCATGTATATATCTTCTGTATG 83931
QY 2109 GCCAGCAATGAATTTCTGCTAGAAATTAAGCCGAGCTGTTGTATATTTAGAGGTGTAAT 2168
DB 83932 GCCAGCAATGAATTTCTGCTAGAAATTAAGCCGAGCTGTTGTATATTTAGAGGTGTAAT 83991
QY 2169 ATTTAGGTTCTGTTGTCAGTCTTTTCTGCGCAATTAACAGTAAGATGTTTAGCAGGTCA 2228
DB 83992 ATTTAGGTTCTGTTGTCAGTCTTTTCTGCGCAATTAACAGTAAGATGTTTAGCAGGTCA 84051
QY 2229 CCTAGTTCGGTTCAGAGAGTCATCATCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2288
DB 84052 CCTAGTTCGGTTCAGAGAGTCATCATCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 84111
QY 2289 TGTTCGGGTTAAGTATGATGAATGCGAGTGTGGCGGCGGTGGTGGTCTCGCCCTGTAA 2348
DB 84112 TGTTCGGGTTAAGTATGATGAATGCGAGTGTGGCGGCGGTGGTGGTCTCGCCCTGTAA 84171
QY 2349 TCTAGCACTTTGGGAGGCGGAGGAGGAGTGTGATCATCCTGAGGTCAAGAGTTCAAGACTAG 2408
DB 84172 TCTAGCACTTTGGGAGGCGGAGGAGGAGTGTGATCATCCTGAGGTCAAGAGTTCAAGACTAG 84231
QY 2409 CTTGGCCCAACATCATGAACCCGCTCTACTTAAATAACAAAAATACAAAAATAGCAGGATGGT 2468
DB 84232 CTTGGCCCAACATCATGAACCCGCTCTACTTAAATAACAAAAATACAAAAATAGCAGGATGGT 84291
QY 2469 GGCACACACCTGTAGTCTCCAGCTACTCGGGAGCCCAACGACGAGAACCGGTTGTACCCA 2528
DB 84292 GGCACACACCTGTAGTCTCCAGCTACTCGGGAGCCCAACGACGAGAACCGGTTGTACCCA 84351
QY 2529 GGAGTGGAGGTTGAGTGGAGCGGAGTTGCAATTCATCTCCACCTCGGCGGCGACAGAG 2588
DB 84352 GGAGTGGAGGTTGAGTGGAGCGGAGTTGCAATTCATCTCCACCTCGGCGGCGACAGAG 84411
QY 2589 CAAGATTCTATC -- AAAAAAAGGAGGAGTGGCAAGTAAAGTTATAGAGGAGGAGGAGGAG 2645
DB 84412 CAAGATTCTATC AAAAAAAGGAGGAGTGGCAAGTAAAGTTATAGAGGAGGAGGAGGAG 84471
QY 2646 CTAGAAGGAATTAAGCGTTGTAGTAAACGCGTGTCTATCTTAAAGTTGAAGAGGAGGAG 2705
DB 84472 CTAGAAGGAATTAAGCGTTGTAGTAAATGCGTGTCTATCTTAAAGTTGAAGAGGAGGAG 84531
QY 2706 ACGAAATCCATTTGTTTAAATTCATCTCAAGGAGGAGGAGAACCCGGGCTGTGTGTGGGT 2765
DB 84532 ACGAAATCCATTTGTTTAAATTCATCTCAAGGAGGAGGAGAACCCGGGCTGTGTGTGGGT 84591
QY 2766 GGTTCGCAATTCCTAGAACGGAATGTGTGGGTATAGAAAAGGAATGAATTAAGCGTTG 2825
DB 84592 GGTTCGCAATTCCTAGAACGGAATGTGTGGGTATAGAAAAGGAATGAATTAAGCGTTG 84651
QY 2826 TTTTTCAAATAGGTCCTTGTAAAGTTATTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2885
DB 84652 TTTTTCAAATAGGTCCTTGTAAAGTTATTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 84711
QY 2886 TTAATAATGATTTGGGAAAAACAATTTGCTTTTGAAGGCTCAGTGACAAACGCAAGGATTACAA 2945
DB 84712 TTAATAATGATTTGGGAAAAACAATTTGCTTTTGAAGGCTCAGTGACAAACGCAAGGATTACAA 84771
QY 2946 CTTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2967
```

```
DB 84772 CTTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 84793

RESULT 7
LOCUS AY219233 2916 bp mRNA linear ROD 20-OCT-2003
DEFINITION Mus musculus novel amplified in breast cancer-1 mRNA, complete cds.
ACCESSION AY219233
VERSION AY219233.1 GI:29648617
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2916)
Beardley, D.I., Kowbel, D., Latexes, T.A., Mannino, J.M., Xin, H.,
Kim, W.J., Collins, C. and Brown, K.D.
Characterization of the novel amplified in breast cancer-1 (NABCL1)
gene product
Exp. Cell Res. 290 (2), 402-413 (2003)
2230202
PUBMED 14567997
REFERENCE 2 (bases 1 to 2916)
AUTHORS Brown, K.
DIRECT SUBMISSION
Submitted (14-JAN-2003) Biochem. and Molec. Biology, LSU Health
Sciences Center, 1901 Perdido Street, New Orleans, LA 70112, USA
LOCATION/Qualifiers
FEATURES
1..2916
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
89..1990
/note="NABCL1; AIBC1; BCAS1"
/codon_start=1
/product="novel amplified in breast cancer-1"
/protein_id="AA088011.1"
/db_xref="GI:29648618"
/translation="MGQMVSPLRPGDQEHDPGADTKVTSNDECNVQNGNPVVLSTR
IQHYEVDLGISSKDNVATSSPKTMEAOAVGDASGNLKGAKTKAPAAASHPTLTL
SRVPGRPGDQTDSSAERFDVPSAAPPENKDPSEHGLPVAAPQAAADPKTPGCP
EAKQOTLPATGLAPSPESQAEAPADQDFGLNRFRLDKGSEAPVNSQPEAKG
SEDEQATEAPAVPNHVSAGSDIVDSQSGQDVTLSVSPGDEVPVPTKEDPQ
VVDTESSIMSFFTLVSPNKTETKDDPEDTKADSVCDGHAAQKMSQKSEQAKS
KKRLDSPRLGLSFKLFRFRDSTNSPTTANLSDKANFTPOETRGKATKSCSP
PPSEPTSEGRDSEKAGPTLLPLGLFWKKSVEKDTLSTGAENAVCESEGVRL
EEVESLQTVDLSEETOPEPTDVKVEESKPKTPLMAFLQMSVRSSEGIPISEEN
VKSSCQTSNVEKTPSPPEPAGTAQKNETSSKDKKSVKKSATENSQKNGKQ
EVREPAFCVQPPPTVEANAMQTGDKTPKSEKRSQSLGGLGFLGPKRMSDAQVQDTPV
SIGPVGKSK"
ORIGIN
Query Match 29.8%; Score 912.8; DB 10; Length 2916;
Best Local Similarity 67.1%; Pred. No. 9.3e-242;
Matches 1564; Conservative 0; Mismatches 567; Indels 201; Gaps 11;
QY 163 CACGACTGGAGACCCAGGACCTGTGCAGGAGCAGCTCCGGGTGACACGAGGGGACTGA 222
DB 1 CAACCGCTGAAGCGCAGGACCTGTGAGCAGNACCACTGTGGAACTGTGGGGGACCCA 60
QY 223 AGATACTCCCAACAGGGCTCAGCAGGAGCAATGGGTAAACCAATAGTGTTCCTCCCAAGA 282
DB 61 AGATACTCCCAACCA--CTTACCAGGAACAATGGAAACCAAGATGAGCGTCCCACTAAGG 118
QY 283 GTTGAACCAACAGAGATGAACCAAGACGACGAGCTTACCAG-----GACACCGG 333
DB 119 CCTGGACACGAGGACGACGACCCAGGAGCAGACTTGCACAGGTGACGTGACAGCAATGAG 178
QY 334 TCTGCTCTGAACGGGGTTCCAGTGTGTGTCGACCCACACACAGTTCAGCACTTAGAGGAA 393
DB 179 TGTGTTCAAGATGGGAACCCAGTTGTGCTGCTCCACTGTGTAAATCAACACTATGAGGA 238
```


QY 394 GTCGACTTGGGAATAAGTGTCAAGACGATTAATGTGGCCACTTTCTTCCCCGAGACAACG 453
Db 239 GTGACTTGGGAATCAGCAGCTCCAAGGATTAATGTGGCTACTTTCTTCCCCAAGACAATG 298
QY 454 GAGATAGTGTGTGGCGATGCCAAGGAAGATCTTGGGAAGAGGCCCAACCCGAG 513
Db 299 GAGGCCCAAGGTGTGGCGATGCCAAGCGGAAGAACTCTTGGGAAGAGGCCCAACCAAG 358
QY 514 GCACGAGCTGTAATCTCGTTTTCTTGATGCTCTCGGCTGTACACGAGGAGTACC 573
Db 359 GCACGAGCTGTAGTCCATCTTTCTTGACATCTCTCGGCTGTACACGAGGAGTCCC 418
QY 574 GGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTTGATGTCAAGTCCAAATAA 633
Db 419 GGGGACCAAGCACGATTCATCGGCTGCATCAGAAAGGTTTGATGTCAAGCCCAAGGCA 478
QY 634 GCTCAGCGAACAAGACCCAAGTGAGAGCTGAGACATTTCCGGTGGCAGCTGGACCGGG 693
Db 479 CGCCTGAGAACAAAGACCCAAGTGAGCACGGGGCACTTCTGTGGCAGCTGCAACCAAGG 538
QY 694 CAGGACACAGATAAAACCCAGGGCACGCCCGGCCCAAGACAAGTCTCTTGC CGCC 753
Db 539 CAGGCTCAGATATAAACCCAGGGTGCCCGAGGCCAAGCAGACACCTTCTTGCACCC 598
QY 754 AGGGATCCCAACGCTTCTCCCACTGAGACAGGGGGAGCAGGAGGAGTCTCCCTCCAAG 813
Db 599 GGCCCATTTGGCACCTCACCACTGAGTCCAG-----GCAGAAGCCCGGCCAG 649
QY 814 CCCAAGGACTCCAGCTTTTGTACAAATCTTCAAGCTGGACAAGGACAGCAAGAAAGTG 873
Db 650 GACAAGGATTTCCGGCTTTTGAACACAGATTTCTTAACTTGGCAAGGAAGAAAGCGCG 709
QY 874 CCAGGTGACAGCAAC---AGGAAGCCAAAGGGCAGAGCATCAAGACAAGTGGATGAG 930
Db 710 CCGGTGAACAGCAGCCCAAGAGAGCGAAAGCTCGGAAGACCCAGAACAGGCCACAGAG 769
QY 931 GTTCTGCTTATCAGGGCAGTCCGATGATGTCCCTGCAGGAAAGGACATAGTTGACGGC 990
Db 770 GCTCTGCGGTGCCAGGAATCCCATCGTGTCTCTGCAAGGGAGGACATAGTCGACAGC 829
QY 991 AAGNAAGAAGACAGAACTTGGAACTGGGATTCGTCTGTCTTCCCTGGGGACCCAGAA 1050
Db 830 GAGCAGA---GAGGACAAGAGCTTGACACTCTGAGTTATTTCTGCTCCCTGGGGATCCAGAG 886
QY 1051 GGACTGGAGACTGCAAAGGAGCATTCACAGCAGCAGCTATAGCAGAGAAATAAATTC 1110
Db 887 GTCCGGGACCAAGAGGAGGACCCGAGGTGTGATACCAAGAGAACAGCAGCTCC 946
QY 1111 ATCATGAGTTCTTTAAACTCTGGTTTCACCTAACAAAGCTGAACAAAAGAGACCCA 1170
Db 947 ATCATGAGCTTCTTCAAGACACTGGTTTCACTTAACAAGACTGAACAAAAGAGACCCA 1006
QY 1171 GAAGAC----- 1178
Db 1007 GAAGACAGGAAGGCAACCAAGGCAGACAGTGTCTGTGATGACACGCTGTGGCCAGAG 1066
QY 1179 ----- 1178
Db 1067 ATGTCGGAGACGAGGCTTAAAGCAAGAAAGCGCCTGGACAGCCCGAGCTAGGACTC 1126
QY 1179 -----GGTGCTGAAAAAGTCAACCCACACTTCAGCTGAC 1212
Db 1127 TCCTTTGAAGAGCTCTTTAGGGCGCAAGGATACTGAGAACTCACCCAGCTTCAGCCAAT 1186
QY 1213 CTTAAGTCAGACAAAGCCAACTTTACATCCAGGAGACCAAGGGCTGGCAAGATTC 1272
Db 1187 CTCAGTCGGAACAAAGCCAACTTCAACCCAGGAGACCCGAGGGAAGCAAGAGGCTACC 1246
QY 1273 AAAGGATCAACCCACTCGGGGCACACAGTCCGTGCAACCCCTGAACTGCGGAAGAA 1332
Db 1247 AAAAGCTGAGCCCCCAGCCG---CACTTCTGNAACCCACAGCGAA---GGGAGAGAC 1300
QY 1333 GGCACCAAGGAGAAATCAGGACCCCACTCTCTGCTCTGGGCAAACTGTTTGGAAAAAG 1392

Db 1301 AGCGCAAGGAGAGGAGGCGGCCCCACTTACTACCGCTGGGAAGTTGTTTGGGAAG 1360
QY 1393 TCAGTTAAAGAGAGACTCAGTCCCCACAGAGTGGCGAGAGAAATGTGTGTGAGTACCA 1452
Db 1361 TCAGTTAAGAGGATACATTTCCACAGGTGCAGAGGAAACCGGTGTGTGAGTACCA 1420
QY 1453 GTAGAGATTATAAGTCCAAAGGAAGTAGAATCAGCCTTACAAAAGTGGAGCTTCAACGAA 1512
Db 1421 GTAGAGCCCTAAGCTTGAAGGATAGAAATCCAGCTTACAAAAGTGGATCTCAGTGA- 1479
QY 1513 GGAGATCTGCAACTGAAACCCACAGAAAGCGAAACTCAAAAGAGAAAGAAACCAAGA 1572
Db 1480 --AGAGACCCAGCCTGAACCCACAGACGTAAAGTCAAAAGAGAAAGCAAAACCCCGAAG 1537
QY 1573 ACCTCTCTGATGGCTTTCTCAGCAAAATGTCTAGTGAAGGGATGAGGGATCACCCAC 1632
Db 1538 ACCCTCTGATGGCTTTCTCAGCAAAATGTCTAGTGAATCGAGCGAAGGATCCCCCGC 1597
QY 1633 TCAGAAATAATAATGGAAAGACTCCAGCTGCCCCAAACATCAGACTCCACAGAAAAGACT 1692
Db 1598 TCGAAGAAAGTAACGTGAAGACTCCAGCTGCCAAACGTCAAACTCCGTGGAAAGACG 1657
QY 1693 ATCAACCCCGCAGAGCTGAAACCAAGGAGACCAAGAGGGTAAAGGGGTCC--- 1749
Db 1658 CCCTCACCCCGCAGAGCCCGCGGGAACAGCTCAGAAAAACAAGAGGACCTCTCTCC 1717
QY 1750 -----TCGAAGGACAAGAACTCAGCAGCCGAGATGAACAAGCAGAAAGAC 1794
Db 1718 TCGAAGACAAGAAATCAGTGGACAGAAGTCTAGCGACTGAGAACACAGACAGAAAGAC 1777
QY 1795 AACAAAGCAGGAAGCAAGAAACAGCCAGCTGCAAGAGAGGCGCACGGTGGACAGAAC 1854
Db 1778 GGCAAAACAGGAAGTCAGAGAGCCGGCGCGGTGTGTGAGCGCCCAAGTGGAGGCAAC 1837
QY 1855 TCACTGCAAGTGGGCAAGCTCCAAAAGAGACTGAGAAAGCGGAGAGTCCCTTGGG 1914
Db 1838 GCATGCAAGTGGGCAAGACCCCAAGAAAGTCTGAGAAGCGGCGACAGTCCCTCGGG 1897
QY 1915 GGCTTCTTTAAAGCCCTGGGACCAAGCGGATGTGATGCTCAAGTGCAGAAACAGACCCA 1974
Db 1898 GGCTTCTGAGGGCTTGGGACCAAGCGGATGTGCGATGCTCAGGTGCAGAACAGGACCG 1957
QY 1975 GTATCCATCGGACAGTGGCAAAACCAAGTAAACAAATCAGCAAGCTTCCACAGGTT 2034
Db 1958 GTCTCCATCGGACCAAGTGGAAATCCAAAGTAAACAAATCACCAGTTCACCACTGGTT 2017
QY 2035 CTCTGCGCACAGATGTGTCTCTTACTTCCATCTCTCTCCCAACACAGCTCCATGTAT 2094
Db 2018 CTCTCTCCAC-----TCCATCTACTGTCCAGACCCACACTGTGTATATATTTTC 2067
QY 2095 ATATTCTTCTGATGGCCAGCAATGAAATCTGCTAGAAATTAAGCCCGAGCTGTGTGA 2154
Db 2068 TTCTTCTTCTGCTGTCAGCAATGAAATCTGCTACACTTCACTGATGATGTTGTA 2127
QY 2155 TATTGAGTGTATTATTACGTCTCTGCTCAGTCTTTTCTGGCAAAATACAGTAAAGAT 2214
Db 2128 GATCAGGTGTATTATTACGTCTCTGCTCAGTCTTTTCTTGCGATTTAAGTGTAAAAAT 2187
QY 2215 GGTTTAGCAGTCACTAGTTGGTTCAGAGAGTCGATGATCACCAGCAGGAAAGGAG 2274
Db 2188 GGCTTAGCAGGTTGACCAAGTTGATCCAAGGAGCTGGTGTGAGTGCAGTGGGGAGGAG 2247
QY 2275 GGAATAGAGGAATCTGTTTCGGTTTAAAGTGAATAATGGCAGTGTGGCGG 2326
Db 2248 TAGGTAGATAAGCAATTTGTGTTAAGTTATGAAAAAATGTCAGGTATCCAG 2299

RESULT 8

AC004501/c

LOCUS

DEFINITION Homo sapiens chromosome 20, BAC clone 121 (LBNL H144), complete sequence.

AC004501

39507 bp

DNA

linear

PRI 20-JUN-1998


```

AC004501
VERSION AC004501.1 GI:2996638
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 39507)
Gray,J.W., Collins,C., Kimmerly,W., Bondoc,M., Cheng,J.,
Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T.,
Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.
Sequencing of human chromosome 20
JOURNAL
TITLE Direct Submission
REFERENCE
AUTHORS Submitted (30-MAR-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
2 (bases 1 to 39507)
Ricke,D.O.
Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
JOURNAL
AUTHORS 3 (bases 1 to 39507)
Gray,J.W., Collins,C., Kimmerly,W., Bondoc,M., Cheng,J.,
Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T.,
Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.
Direct Submission
JOURNAL
TITLE Submitted (30-MAR-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
4 (bases 1 to 39507)
Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
Direct Submission
JOURNAL
TITLE Submitted (30-MAR-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
5 (bases 1 to 39507)
Ricke,D.O.
Direct Submission
JOURNAL
TITLE Submitted (20-JUN-1998) DOE Joint Genome Institute
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES
source
1. 39507
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.2"
/clone="BAC121"
/note="LBNL H144"
misc_feature Complement (1022..1606)
/note="GRAIL 2 excellent exon, frame 1"
repeat_region complement (1981..2192)
/rpt_family="MER20"
repeat_region 2230..2358
/rpt_family="MIR"
repeat_region 2889..2979
/rpt_family="MIR"
repeat_region 4477..7764
/rpt_family="L1"
repeat_region 7749..8061
/rpt_family="Alu"
repeat_region 8035..8054
/note="(A)20"
/rpt_type=tandem
/rpt_unit="a"
repeat_region 8102..8880
/rpt_family="L1"
repeat_region 9213..9313
/rpt_family="MIR"
repeat_region complement (9732..10100)
/rpt_family="THE1b"
repeat_region complement (10253..10394)
/rpt_family="MIR"
repeat_region 13184..13974
/rpt_family="L1"
repeat_region 14234..14580
/rpt_family="Alu"
repeat_region 14544..14565
/note="(A)22"
/rpt_type=tandem
/rpt_unit="a"
repeat_region 14875..15054
/rpt_family="L1"
repeat_region 15932..15964
/note="(TGA)11"
/rpt_type=tandem
/rpt_unit="tga"
repeat_region complement (16272..16455)
/rpt_family="MERS"
repeat_region 17190..17494
/rpt_family="Alu"
repeat_region 17547..17879
/rpt_family="Alu"
repeat_region 18872..19495
/rpt_family="Alu"
repeat_region 19163..19182
/note="(A)20"
/rpt_type=tandem
/rpt_unit="a"
repeat_region 19196..19484
/rpt_family="Alu"
misc_feature complement (20307..20596)
/note="83% identity AC003029"
misc_feature complement (20380..20726)
/note="84% identity B35601"
misc_feature complement (20639..20722)
/note="84% identity zx58h03.r1"
/db_xref="dbEST:AA203298"
repeat_region complement (22197..22354)
/rpt_family="Alu"
repeat_region complement (22435..22514)
/rpt_family="MIR"
repeat_region complement (22575..22996)
/rpt_family="MLT1"
repeat_region complement (24349..24689)
/rpt_family="Alu"
repeat_region 24361..24390
/note="(TTG)10"
/rpt_type=tandem
/rpt_unit="ttg"
repeat_region 24909..25022
/rpt_family="L1"
repeat_region 25103..25246
/rpt_family="MLT1a"
repeat_region 25335..25502
/rpt_family="MLT1"
repeat_region complement (26527..26896)
/rpt_family="THE1B"
repeat_region 27826..27889
/rpt_family="L1MD1"
repeat_region complement (27900..28255)
/rpt_family="THE1"
repeat_region 28248..29228
/rpt_family="L1"
repeat_region complement (29552..29682)
/rpt_family="MIR"
repeat_region complement (29955..30101)
/rpt_family="MIR"
repeat_region 30224..30576
/rpt_family="MLT1"
misc_feature complement (30651..30788)
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 31228..31247
/note="(T)20"
/rpt_type=tandem
/rpt_unit="t"

```



```
* 1 1374: contig of 1374 bp in length
* 1375 1474: gap of 100 bp
* 1475 3271: contig of 1797 bp in length
* 3272 3371: gap of 100 bp
* 3372 5701: contig of 2330 bp in length
* 5702 5801: gap of 100 bp
* 5802 8678: contig of 2877 bp in length
* 8679 8778: gap of 100 bp
* 8779 11405: contig of 2627 bp in length
* 11406 11505: gap of 100 bp
* 11506 15628: contig of 4123 bp in length
* 15629 15728: gap of 100 bp
* 15729 18099: contig of 2371 bp in length
* 18100 18199: gap of 100 bp
* 18200 21769: contig of 3570 bp in length
* 21770 21869: gap of 100 bp
* 21870 25732: contig of 3863 bp in length
* 25733 31618: contig of 5786 bp in length
* 31619 31718: gap of 100 bp
* 31719 36866: contig of 5148 bp in length
* 36867 36966: gap of 100 bp
* 36967 46305: contig of 9339 bp in length
* 46306 46405: gap of 100 bp
* 46406 58028: contig of 11623 bp in length
* 58029 58128: gap of 100 bp
* 58129 72476: contig of 14348 bp in length
* 72477 72576: gap of 100 bp
* 72577 88607: contig of 16031 bp in length
* 88608 88707: gap of 100 bp
* 88708 102833: contig of 14126 bp in length
* 102834 127147: contig of 24214 bp in length
* 127148 127247: gap of 100 bp
* 127248 150561: contig of 23314 bp in length
* 150562 150661: gap of 100 bp
* 150662 181313: contig of 30652 bp in length.
FEATURES
    source
    1..181313
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="20"
        /map="20"
        /clone="RP11-469D15"
        /clone_lib="RPC1-11 Human Male BAC"
    1..1374
        /note="assembly_fragment"
    1475..3271
        /note="assembly_fragment"
    3372..5701
        /note="assembly_fragment"
    5802..8678
        /note="assembly_fragment"
    8779..11405
        /note="assembly_fragment"
    11506..15628
        /note="assembly_fragment"
    15729..18099
        /note="assembly_fragment"
    clone_end:17
        vector_side:left"
    18200..21769
        /note="assembly_fragment"
    21870..25732
        /note="assembly_fragment"
    25833..31618
        /note="assembly_fragment"
    31719..36866
        /note="assembly_fragment"
    36967..46305
        /note="assembly_fragment"
    46406..58028
```

```
misc_feature
58129..72476
/note="assembly_fragment"
misc_feature
72577..88607
/note="assembly_fragment"
misc_feature
88708..102833
/note="assembly_fragment"
misc_feature
102934..127147
/note="assembly_fragment"
misc_feature
127248..150561
/note="assembly_fragment"
misc_feature
150662..181313
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
ORIGIN
Query Match 19.0%; Score 583; DB 2; Length 181313;
Best Local Similarity 99.2%; Pred. No. 6.5e-150; Indels 0; Gaps 0;
Matches 586; Conservative 0; Mismatches 5;
QY 393 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGGAGACAAC 452
Db 137836 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGGAGACAAC 137777
QY 453 GGAGATAAGTCTGTTGGGATGCCAAGCGAAGAAATCTTGGGAAAGAGGCCAACCCGA 512
Db 137776 GGAGATAAGTCTGTTGGGATGCCAAGCGAAGAAATCTTGGGAAAGAGGCCAACCCGA 137717
QY 513 GGCACCAAGCTCTAAATCTCGTTTTTCTTGATGCTCTCTCGGCTGTACAGGACGTAC 572
Db 137716 GGCACCAAGCTCTAAATCTCGTTTTTCTTGATGCTCTCTCGGCTGTACAGGACGTAC 137657
QY 573 CGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTTGTATGTGAGTCCAATAA 632
Db 137656 CGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTTGTATGTGAGTCCAATAA 137597
QY 633 AGTCCAGCGAACAAGACCCCAAGTGAGAGTGACACTTCCGTTGGCAGTGGACCCGG 692
Db 137596 AGTCCAGCGAACAAGACCCCAAGTGAGAGTGACACTTCCGTTGGCAGTGGACCCGG 137537
QY 693 GCAGGACACAGATAAACCAGGCGACGCCCGGCCCAAGCAAGGTCTCTCTGCGGC 752
Db 137536 GCAGGACACAGATAAACCAGGCGACGCCCGGCCCAAGCAAGGTCTCTCTGCGGC 137477
QY 753 CAGGAGTCCCAAGCTTCTCCCACTGAGACAGGGGAGCAGGAGAGAGTCCCTCCAA 812
Db 137476 CAGGAGTCCCAAGCTTCTCCCACTGAGACAGGGGAGCAGGAGAGAGTCCCTCCAA 137417
QY 813 GCCCAAGACTCCAGCTTTTGTGACAAATTTTCAAGCTGGACAGGACAGGAAAAGGT 872
Db 137416 GCCCAAGACTCCAGCTTTTGTGACAAATTTTCAAGCTGGACAGGACAGGAAAAGGT 137357
QY 873 GCCAGGTGACAGCCAAAGGAGCCAGAGGGGAGAGCAGTCAAGCAGGTGGATGAGGT 932
Db 137356 GCCAGGTGACAGCCAAAGGAGCCAGAGGGGAGAGCAGTCAAGCAGGTGGATGAGGT 137297
QY 933 TCCTGGCTTATCAGGCGAGTCCGATGATGTCCCTGAGGAGAGGACATAGT 983
Db 137296 TCCTGGCTTATCAGGCGAGTCCGATGATGTCCCTGAGGAGAGGAGT 137246
```

RESULT 10
AB029897

LOCUS AB029897 1885 bp mRNA linear ROD 02-JUL-2002
DEFINITION Rattus norvegicus mRNA for band83, complete cds.
ACCESSION AB029897
VERSION AB029897.1 GI:21668479
KEYWORDS band83.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;


```
RESULT 11
AX150111
LOCUS AX150111 488 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 86 from Patent WO0136685.
ACCESSION AX150111
VERSION AX150111.1 GI:14348139
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Kroes, R.A., Moskal, J.R. and Yamamoto, H.
TITLE Differential gene expression in cancer
JOURNAL Patent: WO 0136685-A 86 25-MAY-2001;
NYXIS Neurotherapies, Inc. (US)
FEATURES
Location/Qualifiers
source 1..488
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 15.3%; Score 470.4; DB 6; Length 488;
Best Local Similarity 99.8%; Pred. No. 5.6e-119;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2142 CCGAGCTGTGATATTCAGGCTGATATTTACGCTCTCGTCCAGTCTTTCTGGCAAA 2201
Db 15 CCGAGCTGTGATATTCAGGCTGATATTTACGCTCTCGTCCAGTCTTTCTGGCAAA 74
QY 2202 TAACAGTAAAGATGTTTACGAGTCACTAGTTGGTCCAGAGTGCATGATCACC 2261
Db 75 TAACAGTAAAGATGTTTACGAGTCACTAGTTGGTCCAGAGTGCATGATCACC 134
QY 2262 GCAGAAAGGGAGGAATAGAGGATGTTCCGGTTAAGTATGATCAAAATGGCAGTGTG 2321
Db 135 GCAGAAAGGGAGGAATAGAGGATGTTCCGGTTAAGTATGATCAAAATGGCAGTGTG 194
QY 2322 GCCGGCGTGTGGCTCTCGCTGTAATCTCAGACACTTTGGAGGCCGAGGAGTGGAT 2381
Db 195 GCCGGCGTGTGGCTCTCGCTGTAATCTCAGACACTTTGGAGGCCGAGGAGTGGAT 254
QY 2382 CACTTGAGTCCAGGATTCAGACTAGCTGGCCCAACATCATGAAACCCCGTCTCTACTA 2441
Db 255 CACTTGAGTCCAGGATTCAGACTAGCTGGCCCAACATCATGAAACCCCGTCTCTACTA 314
QY 2442 AAAATACAAAATAGCCAGGATGTCGCACACACCTGTAGTCCAGCTACTCGGGAGC 2501
Db 315 AAAATACAAAATAGCCAGGATGTCGCACACACCTGTAGTCCAGCTACTCGGGAGC 374
QY 2502 CCAAGCCAGGAAACCGCTTGACCCAGAGGTGGAGGTGCAGTGAGCCGAAGTTGCAC 2561
Db 375 CCAAGCCAGGAAACCGCTTGACCCAGAGGTGGAGGTGCAGTGAGCCGAAGTTGCAC 434
QY 2562 CATTGCATCCACCTCGGCGCAGAGCAAGATTCTATCAAAAAAAGG 2613
Db 435 CATTGCATCCACCTCGGCGCAGAGCAAGATTCTATCAAAAAAAGG 486
RESULT 12
AL935134/c
LOCUS AL935134 105616 bp DNA linear ROD 01-JUL-2003
DEFINITION Mouse DNA sequence from clone RP23-321D1 on chromosome 2, complete
sequence.
ACCESSION AL935134
VERSION AL935134.10 GI:32400093
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

1 (bases 1 to 105616)

Humphries, M.

Direct Submission

Submitted (01-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jul 1, 2003 this sequence version replaced gi:2988596.

Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun map have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information
on the WormPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-321D1 is

from the RP23-23 Mouse BAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES
source

1..105616

Location/Qualifiers

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="2"

/clone="RP23-321D1"

/clone_lib="RPCI-23"

ORIGIN

Query Match 9.9%; Score 303.6; DB 10; Length 105616;
Best Local Similarity 72.1%; Pred. No. 2.4e-72;
Matches 428; Conservative 0; Mismatches 154; Indels 12; Gaps 2;

QY 393 AGTCGACTTGGGATTAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGGAGACAAC 452

Db 9080 AGTCGACTTGGGATTAAGTGTCAAGACGGATAATGTGGCCACTTCTTCTTCCCGGAGACAAC 9021

QY 453 GGAGATTAAGTGTCTGTTCGGATGCAACGGAAAGAAATCTTCGGAAAGAGCCAAACCCGA 512

Db 9020 GGAGCCAGGCTGTGGCGATGCGAGGAAAGAAATCTTCGGAAAGAGCCAAACCCGA 8961

QY 513 GGCAACGAGCTGCTAAATCTCGTTTTTTTCTTGATGCTCTCTCGGCTGTACAGGACGTAC 572

Db 8960 GGCAACGAGCTGCTAGATCCCAATTTTCTTGACACTCTCTCGGCTGTACAGGACGTAC 8901

QY 573 CGGAGACGAGCCGACGATTCATCCCTTGGATCAGTGAAGCTTGTATGTCAGCTCAATAA 632

Db 8900 CGGAGACGAGCCGACGATTCATCCCTTGGATCAGTGAAGCTTGTATGTCAGCTCAATAA 8841

QY 633 AGCTCCAGGCAACCAAGACCAAGTGAAGCTGACACTTCCGCTGGCAGTGGACCGG 692

Db 8840 AGCTCCAGGCAACCAAGACCAAGTGAAGCTGACACTTCCGCTGGCAGTGGACCGG 8781

Qy 693 GCAGGACACAGATAAAACCCAGGCGACGCCCGGCCCAAGACAAGGTCCTCTCTGCGGC 752
 |||||
 Db 8780 CGAGGCTCCAGATAAAACCCAGGCGTCCCGAGGCCAAGCAGCAGACCTCTCTGCGCAC 8721
 |||||
 Qy 753 CAGGATCCCAAGCTTCTCCACCTGACAGAGGGGAGCAGGAGAGAGCTCCCTCCAA 812
 |||||
 Db 8720 CGGCGCATTTGGCACCCTCACCACCTAGTCCAGGC-----AGAAGCCCGGCCCA 8670
 |||||
 Qy 813 GCCCAAGACTCCAGCTTTTGTGCAATTTTCAAGCTGCAACAGGCAAGGCAAGAAAGGT 872
 |||||
 Db 8669 GGACAAGGATTTGCGCTTTTGAACAGATCTTTAAACTTGGACAAAGAGAGAAAGCGC 8610
 |||||
 Qy 873 GCCAGGTGACAGCCAAC---AGGAAGCCAAGGCGCAGAGCATCAAGACRAAGTGGATGA 929
 |||||
 Db 8609 GCCGGTGAACAGCAGCCCAAGGAAGGCTCGGAAGACCCAGACAGGCCACAGA 8550
 |||||
 Qy 930 GGTCTCTGGCTTATCAGGGCAGTCCGATGATGTCCTGTCAGGGAAGGACATAGT 983
 |||||
 Db 8549 GGCTCTCTGCGTGCAGGGAATCCCATGCTCTCTGTCAGGGGAGGTAAAGTGT 8496
 |||||

RESULT 13

AC126535
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-408J13, *** SEQUENCING IN PROGRESS
 *** 3 unordered pieces.

AC126535

AC126535.3 GI:23196178

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 168876)

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsebrook, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Avogri, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregregis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovat, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensuhera, L., Loulsegod, H., Lozado, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mawney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, I.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokelemeh, O., Okwuon, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,

Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajes, D.,
 Shedd, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Swatek, A., Tabor, P., Taylor, C.,
 Taylor, I., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 168876)

Worley, K.C.

Direct Submission

Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 168876)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (19-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the
 sequence may extend beyond the ends of the clone and there may be
 contigs that consist entirely of whole genome shotgun sequence
 reads. Both end sequences and whole genome shotgun sequence only
 contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GyFD

Center clone name: CH230-408J13

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 151329 bases at least Q40

Consensus quality: 153110 bases at least Q30

Consensus quality: 154290 bases at least Q20

Estimated insert size: 162792; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces

* is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will
 * be preserved.

* 1 12922: contig of 12922 bp in length

* 12923 13022: gap of unknown length

* 13023 155459: contig of 142437 bp in length

* 155460 155559: gap of unknown length

* 155560 168876: contig of 13317 bp in length.

Location/Qualifiers

1..168876

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clones="CH230-408J13"

11757..12922

/note="wgs_contig"

FEATURES
source

misc_feature

```

misc_feature      13023..14162
                  /note="wgs_contig"
misc_feature      15743..16938
                  /note="wgs_contig"
ORIGIN
Query Match      9.9%; Score 303.6; DB 2; Length 168876;
Best Local Similarity 72.1%; Pred No. 2.6e-72;
Matches 428; Conservative 0; Mismatches 154; Indels 12; Gaps 2;
QY 393 AGTCAGCTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGAGACACAC 452
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 58999 AGTCGATTGGGAATCAACAGCTTGAAGGATAATGTGGCCGCTTCTTCCCGAGACACAT 59058
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 453 GGAGATAAGTGTCTGTGGGATGCCAACGGAAGAATCTTGGGAAGAGGCCAACCCGGA 512
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59059 GGAGGCTCAGCGTGTGGGTGATGCCGCGGAAGAATCTTGGGAAGAGATCCAGCCCAA 59118
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 513 GGCACGAGCTGCTAAATCTCGTTTTTTCTTGATGCTCTCTGGCCCTGTACAGGACGTAC 572
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59119 AGCAGAGCTGTAGTCCATTTTCTTGACACTCTCTCGGCTGTACAGGACGTCC 59178
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 573 CGGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTTGTGATGCTCCCAATAA 632
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59179 CGGGACCAAGGCACAGATTTCATCGCTGCATCGGGAGGCTTGTGATGCTAGCCAGGCAC 59238
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 633 AGCTCCAGCAACAAAGACCCCAAGTGAGAGCTGACATTCCTGGTGGCAGCTGGACCGGG 692
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59239 GCGCGCTGAGAACAAAGACCCGAGTGAGCAGCGGGCACTTCCGGTGGCAGCTGCACACAGG 59298
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 693 CGAGGACACAGATAAAACCCAGGCGCAGCCCGCCCAAGACAGTCTCTCTGCGCG 752
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59299 GCGGCTCAGATAAAACCCAGGCTGACTGAGGCCAAGCAGACCTCTCCCTGGCCAC 59358
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 753 CAGGATCCCAAGCTTCTCCCACTGAGACAGGGGGAGCAGGAGAGAGTCTCCCTCCAA 812
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59359 CAGCCCACTGGCACCTTCCACCACTGAGTCCACGCGC-----AGAAGCCCGCTCCAG 59409
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 813 GCCCAAGACTCCAGCTTTTGTGACAAATCTTCAAGCTGACAGGAGGACAGGAAAAGGT 872
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59410 GCCCAAGGATTTTCACTTTTGAACAGACTCTTTAACTGGCAAGGGAAGAGAAAGCGC 59469
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 873 GCCAGGTGACAGCCAAAC---AGGAGCCAGAGGCGCAGCATCAGACAGGTGTGATGA 929
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59470 GCGGCTCAGACGCGCGCTGGAAGAGGGAAGGGTCCAAAAGCCAGGAGCGCCATAGA 59529
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 930 GTTCTCTGCTTATCAGCGCACTCCGATGATGCTCCCTGCGAGGAGGACATAGT 983
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59530 GGTCTCTGCTGTACAGGGAATCCTCTGCTGCTGCTGCGAGGAGGTAAGTGT 59583
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
AC094145
LOCUS              233233 bp      DNA      linear      HTG 09-MAY-2003
DEFINITION        Rattus norvegicus clone CH230-2L13, WORKING DRAFT SEQUENCE, 6
                  unordered pieces.
ACCESSION         AC094145
VERSION           AC094145.6 GI:30467887
KEYWORDS          HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE            Rattus norvegicus (Norway rat)
ORGANISM          Rattus norvegicus
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                  Rattus.
REFERENCE
  1 (bases 1 to 233233)
  Murzyn D, Marie, Metaker, M Lee., Abramzon, S., Adams, C., Alder, J.,
  Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
  Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
  Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
  Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
  Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
  Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
  Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

```

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Deigado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorjis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladik, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorusu, H., Loulseg, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfanckoch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vega, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederstock, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 233233)
Worley, K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 233233)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24942806.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

```
----- Project Information
Center project name: GABR
Center clone name: CH230-2L13
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 207295 bases at least Q40
Consensus quality: 210461 bases at least Q30
Consensus quality: 212871 bases at least Q20
Estimated insert size: 217628; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 6 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.
  1 89147: contig of 89147 bp in length
  * 89148 89247: gap of unknown length
  * 89248 99513: contig of 10266 bp in length
  * 99514 99613: gap of unknown length
  * 99614 113910: contig of 14297 bp in length
  * 113911 114010: gap of unknown length
  * 114011 230487: contig of 116477 bp in length
  * 230488 230587: gap of unknown length
  * 230588 232001: contig of 1414 bp in length
  * 232002 232101: gap of unknown length
  * 232102 233233: contig of 1132 bp in length.
  Location/Qualifiers
    1..233233
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-2L13"
  misc_feature
    1..1412
      /note="wgs_contig"
ORIGIN
Query Match          9.9%; Score 303.6; DB 2; Length 233233;
Best Local Similarity 72.1%; Pred. No. 2.7e-72;
Matches 428; Conservative 0; Mismatches 154; Indels 12; Gaps 2;

Qy 393 AGTCAGCTTGGGATAAGTGTCAAGACGGATAATGTGCGCACTTCTCCCGGAGACAAAC 452
Db 215357 AGTCGATTGGGAATCAACAGCTTGAAGGATAATGTGCCCGCTTCTTCCCGGAGACAAT 215416

Qy 453 GGAGATAAGTGTGTTGGGATGCCAAGGAAAGAAATCTTGGGAAAGAGGCCAAACCCGA 512
Db 215417 GGAGCTCACGCTGTGGCTGTATGCCAGCGGAAGAAATCTTGGGAAAGATCAAGCCCA 215476

Qy 513 GGCACCACTGCTAAATCTGTTTTTCTTGATGCTCTCTCGGCTGTACAGGACGTPAC 572
Db 215477 AGCACCACTGCTAGTCCCAATTTTCTTGACACTCTCTCGGCTGTACAGGACGTPCC 215536

Qy 573 CGGAGACCAAGCCGAGATTATCCCTTGGATCAGTGAAGCTTGATGTGCTCAATAA 632
Db 215537 CGGGACCAAGGCACAGATTATCCGCTGCACTCGGGGAGGTTGATGTCAAGCCAGGCAC 215596

Qy 633 AGCTCCAGCAACAAAGACCAAGTGAGAGCTGGACACTTCGGTGGGAGCTGGACCGGG 692
Db 215597 GCGGCTTGAGAAACAAGACCCGAGTGAGCACGGGCACTTTCGGTGGGAGCTGCACACAG 215656

Qy 693 GCAGGACACAGATAAAACCCAGGCGACGCCCGCCCAAGACAAAGTCTCTCTGCGCG 752
Db 215657 GCGGCTCCAGATAAAACCCAGGCGTCACTGAGGCCAAGCAGCAGACCTTCCCTGCCAC 215716

Qy 753 CAGGATCCACGCTTCTCCACCTGACAGCGGGGAGCAGGAGAGAGCTCCCTCCAA 812
Db 215717 CAGCCCACTGGCACCTTCAACCACTGAGTCCAGGC-----AGAAAGCCCGTCCAG 215767

Qy 813 GCCCAGGACTCCAGCTTTTTCACAAATTTTTCAGCTGGACAGGACAGGAAAGGT 872
Db 215768 GCCCAGGATTTCAGCTTTTTCAGACAGACTTTTAACTTCGACAAAGGAGAGAAAGCGC 215827

Qy 873 GCACAGTTCAGCAGCAAC---AGGAAGCCAGAGGGCAGAGCATCAAGACAAAGTGGATGA 929
Db 215828 GCGGCTGACAGCGCGCTCGAAGAAAGGAGGTCCAAAGACCAAGACGGGCCATAGA 215887

Qy 930 GGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCTGCGAGGGAAGCACATAGT 983
Db 215888 GGCTCTCTGCTACCAAGGAATCTCTCATGTGTCTCTGCGAGGAAGGTAAGTGT 215941

RESULT 15
AC084066 235411 bp DNA linear HTG 12-OCT-2000
LOCUS Mus musculus clone RP23-321D1, *** SEQUENCING IN PROGRESS ***, 29
DEFINITION unorderd pieces.
AC084066
VERSION AC084066.1 GI:10799415
KEYWORDS HTG; HTGS PHASE1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 235411)
            DOE Joint Genome Institute.
AUTHORS Sequencing of Mouse
TITLE Unpublished
JOURNAL 2 (bases 1 to 235411)
            DOE Joint Genome Institute.
REFERENCE DOE Joint Submission
AUTHORS Submitted (12-OCT-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
TITLE Direct Submission
JOURNAL
COMMENT
            -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----
Project Information
Center Project Name: 2351294
Center clone name: RPCI-23_321D1
-----
Summary Statistics
Consensus quality: 214207 bases at least Q40
Consensus quality: 223053 bases at least Q30
Consensus quality: 225208 bases at least Q20
Estimated insert size: 200000; pulse field gel estimation
Estimated insert size: 232611; sum-of-contigs estimation
Quality coverage: 9.94 in Q20 bases; pulse field gel estimation
Quality coverage: 11.56 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
  consists of 29 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.
  1 1210: contig of 1210 bp in length
  * 1211 1310: gap of unknown length
  * 1311 3240: contig of 1930 bp in length
  * 3241 3340: gap of unknown length
  * 3341 4504: contig of 1164 bp in length
  * 4505 4604: gap of unknown length
  * 4605 6172: contig of 1568 bp in length
  * 6173 7354: contig of 1082 bp in length
  * 7355 7454: gap of unknown length
  * 7455 8625: contig of 1171 bp in length
  * 8626 8726 10114: contig of 1389 bp in length
```



```
* 10115 10214: gap of unknown length
* 10215 12091: contig of 1877 bp in length
* 12092 12191: gap of unknown length
* 12192 14113: contig of 1922 bp in length
* 14114 14213: gap of unknown length
* 14214 15286: contig of 1073 bp in length
* 15287 15386: gap of unknown length
* 15387 17839: contig of 2453 bp in length
* 17840 17939: gap of unknown length
* 17940 20266: contig of 2327 bp in length
* 20267 20366: gap of unknown length
* 20367 22001: contig of 1635 bp in length
* 22002 22101: gap of unknown length
* 22102 24307: contig of 2206 bp in length
* 24308 24407: gap of unknown length
* 24408 26405: contig of 1998 bp in length
* 26406 26505: gap of unknown length
* 26506 29691: contig of 3186 bp in length
* 29692 29791: gap of unknown length
* 29792 34705: contig of 4914 bp in length
* 34706 34805: gap of unknown length
* 34806 39749: contig of 4944 bp in length
* 39750 39849: gap of unknown length
* 39850 45296: contig of 5447 bp in length
* 45297 45396: gap of unknown length
* 45397 51476: contig of 6080 bp in length
* 51477 51576: gap of unknown length
* 51577 59008: contig of 7432 bp in length
* 59009 59108: gap of unknown length
* 59109 66218: contig of 7110 bp in length
* 66219 66318: gap of unknown length
* 66319 76778: contig of 10460 bp in length
* 76779 76878: gap of unknown length
* 76879 80098: contig of 21220 bp in length
* 80099 98198: gap of unknown length
* 98199 113987: contig of 15789 bp in length
* 113988 114087: gap of unknown length
* 114088 132115: contig of 18028 bp in length
* 132116 160184: contig of 27969 bp in length
* 160185 160284: gap of unknown length
* 160285 186698: contig of 26414 bp in length
* 186699 186798: gap of unknown length
* 186799 235411: contig of 48613 bp in length.

FEATURES
      source
      1..235411
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10090"
      /clone="RP23-321D1"
      /clone_lib="RPC1 mouse BAC library 23"
```

ORIGIN

```
Query Match      9.9%; Score 303.6; DB 2; Length 235411;
Best Local Similarity 72.1%; Pred. No. 2.7e-72;
Matches 428; Conservative 0; Mismatches 154; Indels 12; Gaps 2;

Qy      393 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTCTTCCCGGAGACAAC 452
Db      84020 AGTCGACTTGGGAATCAGCAGCTCCAGAGTAATGTGGCTACTTCTTCCCGGAGACAAT 84079

Qy      453 GGAGATAAGTGTGTTCGGATGCCAACCGGAAGAATCTTGGGAAGAGGCCAACCCGA 512
Db      84080 GGAGGCCCGAGGTGTGGGCGATGCCAGCGGAAGAATCTTGGGAAGAAGCCAGACCA 84139

Qy      513 GGCACCACTGCTAAATCTCGTTTCTTCTGATGTCTCTCGGCGTGTACGAGGACGTAC 572
Db      84140 GGCACCACTGCTAGATCCCAATTTTCTTGACACTCTCTCGGCGTGTACGAGGACGTCC 84199

Qy      573 CGGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTTGTATGTGAGTCTCAATAA 632
Db      84200 CGGGACCAAGCCAGCGATTATCGGCTGCATCAGGAGGTTGTATGTGAGGCGCCAGCGC 84259
```

```
Qy      633 AGCTCCAGCGAACAAAGACCCCAAGTGAAGCTGGACACTTCCGGTGGCAGCTGGACCGG 692
Db      84260 AGCGCTGAGAAACAAAGACCCCAAGTGAAGTGAACCGGGGCACTTCCGGTGGCAGCTGGACCCAGG 84319

Qy      693 GCAGGACACAGATAAAACCCCGAGGCGACGCGCCCGGCCCAAGACAAGGTCTCTTCTGCGGC 752
Db      84320 GCAGGCTCCAGATAAACCCTCCAGGGTGCCTCCGAGGCGCAAGCAGCAGACCTCTCCCTGCCAC 84379

Qy      753 CAGGGATCCCAAGCTTCTTCCCACTGAGACAGGGGGAGCAGGAGGAGAAAGCTCCCTCCAA 812
Db      84380 CGGCCCCATTTGGCACCCTCACCACCTGAGTCCAGGC-----AGAAGCCCCCGGCCCA 84430

Qy      813 GCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGACAGGACAGGAAAGGT 872
Db      84431 GGACAAGGATTTTCGGCTTTTGTGAAACAGATTCTTTAACTGGACAAAGGAAGAAAGCGC 84490

Qy      873 GCCAGGTGACAGCCAAC---AGGAAGCCCAAGAGGCGAGAGCATCAAGACAAGGTGGATGA 929
Db      84491 GCCGTGAAACAGCAGCCCAAGAAAGCGAAAGGCTCGGAAGACCCAGAACAGGCCACAGA 84550

Qy      930 GGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCTTCTGACGGGAGGACATAGT 983
Db      84551 GGCTCCTGCCGTGCCAGGGAATCCCATGGTGTCTCTGACGGGAGGTAAGTGT 84604
```

Search completed: July 4, 2005, 01:47:00
Job time : 13007 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 05:45:02 ; Search time 1566.51 Seconds

(without alignments)
11586.223 Million cell updates/sec

Title: US-08-731-499-12
Perfect score: 3066
Sequence: 1 GGACACGCTATGACCATGA.....GGTACCAATTGCGCCTATA 3066

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:.*
1: Geneseq_1980s:.*
2: Geneseq_1990s:.*
3: Geneseq_2000s:.*
4: Geneseq_2001as:.*
5: Geneseq_2001bs:.*
6: Geneseq_2002as:.*
7: Geneseq_2002bs:.*
8: Geneseq_2003as:.*
9: Geneseq_2003bs:.*
10: Geneseq_2003cs:.*
11: Geneseq_2003ds:.*
12: Geneseq_2004as:.*
13: Geneseq_2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 3066 | 100.0 | 3066 | 2 | AAV09025 Homo sapi |
| 2 | 2828.8 | 92.3 | 2418 | 12 | Adq23347 Human sof |
| 3 | 2813 | 91.7 | 2813 | 5 | As81115 DNA encod |
| 4 | 2813 | 91.7 | 2813 | 6 | Ad38092 Human BCA |
| 5 | 2813 | 91.7 | 2813 | 12 | AdJ75127 Marker ge |
| 6 | 2813 | 91.7 | 2813 | 12 | AdQ18993 Human sof |
| 7 | 2813 | 91.7 | 2813 | 13 | AdR25076 Breast ca |
| 8 | 1923.4 | 62.9 | 2020 | 8 | Acc90605 Human CGD |
| 9 | 1567 | 51.1 | 2105 | 4 | As812521 Gene #14 |
| 10 | 1567 | 51.1 | 2105 | 10 | Abq77258 Human ten |
| 11 | 1016.4 | 33.2 | 9889 | 5 | AbA17779 Human ner |
| 12 | 584 | 19.0 | 1504 | 8 | Abz36099 Human sec |
| 13 | 582.6 | 19.0 | 644 | 12 | Ach89479 Human gen |
| 14 | 470.4 | 15.3 | 488 | 4 | Ah50757 Human tum |
| 15 | 332.8 | 10.9 | 576 | 12 | Ach70719 Human gen |
| 16 | 328 | 10.7 | 372 | 10 | Adf85762 Human ade |
| 17 | 313 | 10.2 | 382 | 4 | AI80677 Human pol |
| 18 | 297 | 9.7 | 506 | 12 | Ach75758 Human gen |
| 19 | 289.2 | 9.4 | 540 | 13 | AdQ53268 Novel can |
| 20 | 265 | 8.6 | 265 | 4 | AA152231 Probe #20 |

| | | | | | |
|----|-------|-----|--------|----|---------------------|
| 21 | 265 | 8.6 | 265 | 4 | AAK46334 Human bon |
| 22 | 265 | 8.6 | 265 | 4 | AAK20263 Human bra |
| 23 | 265 | 8.6 | 265 | 4 | ABs46058 Human liv |
| 24 | 265 | 8.6 | 265 | 6 | ABs20652 Human gen |
| 25 | 260.4 | 8.5 | 495 | 4 | AAI39177 Probe #78 |
| 26 | 260.4 | 8.5 | 495 | 4 | AAK33393 Human bon |
| 27 | 260.4 | 8.5 | 495 | 4 | AAK07595 Human bra |
| 28 | 260.4 | 8.5 | 495 | 4 | ABs33149 Human liv |
| 29 | 260.4 | 8.5 | 495 | 6 | ABs08233 Human gen |
| 30 | 247 | 8.1 | 349 | 2 | AAV88411 EST clone |
| 31 | 242.6 | 7.9 | 6096 | 6 | ABK92513 Human pro |
| 32 | 240.8 | 7.9 | 214520 | 10 | ADL13471 Osteoartrh |
| 33 | 238 | 7.8 | 297 | 3 | AAK10324 Human sec |
| 34 | 236 | 7.7 | 181684 | 11 | ACN44374 Human gen |
| 35 | 233.8 | 7.6 | 68355 | 8 | ACF62737 Cancer ba |
| 36 | 233.8 | 7.6 | 68355 | 8 | ADB20852 MRP1 base |
| 37 | 233.8 | 7.6 | 68355 | 10 | ADB87941 Human UGT |
| 38 | 233.8 | 7.6 | 68355 | 10 | ADB96924 Human MDR |
| 39 | 233.8 | 7.6 | 68355 | 10 | ADB92115 Human MDR |
| 40 | 233.8 | 7.6 | 186591 | 8 | ACF62750 Cancer ba |
| 41 | 233.8 | 7.6 | 186591 | 8 | ADB20869 MRP1 base |
| 42 | 233.8 | 7.6 | 186591 | 10 | ADB87958 Human UGT |
| 43 | 233.8 | 7.6 | 186591 | 10 | ADB96941 Human MDR |
| 44 | 233.8 | 7.6 | 186591 | 10 | ADB92132 Human MDR |
| 45 | 233.8 | 7.6 | 208648 | 8 | ACF62735 Cancer ba |

ALIGNMENTS

RESULT 1
AAV09025
ID AAV09025 standard; cDNA; 3066 BP.
XX
AC AAV09025;
XX
DT 21-JUL-1998 (first entry)
XX
DE Homo sapiens 20ql3 amplicon lbl transcript.
XX
KW 20ql3 amplicon; chromosome 20; tumour; detection;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentation;
KW Leber's congenital amaurosis; ds.
XX
OS Homo sapiens.
XX
PN WO9802539-A1.
XX
PD 22-JAN-1998.
XX
PF 15-JUL-1997; 97WO-US012343.
XX
PR 15-JUL-1996; 96US-00680395.
PR 16-OCT-1996; 96US-00731499.
XX
PA 17-JAN-1997; 97US-00785532.
XX
PI (REGC) UNIV CALIFORNIA.
XX
PI Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX
WPI; 1998-110587/10.
XX
DR New sequences from the 20ql3 amplicon - used for detecting chromosomal
XX abnormalities, particularly tumours, and for developing products for
XX treating diseases.
XX
PS Claim 1; Page 69-70; 91pp; English.
XX
CC The sequence is that of a cDNA sequence lbl, which was isolated from the
XX 20ql3 amplicon, it is overexpressed in numerous breast cancer cell lines
XX and some primary tumours. It can be used as a probe for the detection of
XX chromosomal abnormalities at 20ql3. It and other sequences isolated from

CC the 20q13 amplicon are consistently amplified in primary tumours. These
CC sequences are useful as probes or as probe targets for monitoring the
CC relative copy number of corresponding sequences from a biological sample
CC such as tumour cells. The sequences can also be used in therapeutic
CC applications for modulating the expression of the endogenous gene or the
CC activity of the gene product. Examples of therapeutic approaches include
CC antisense inhibition of gene expression, gene therapy, and monoclonal
CC antibodies that specifically bind the gene products. The products can
CC also be used in the treatment of other diseases, e.g. age-related macular
CC degeneration, Leber's congenital amaurosis and retinitis pigmentation
XX
SQ Sequence 3066 BP; 944 A; 741 C; 818 G; 563 T; 0 U; 0 Other;

Query Match 100.0%; Score 3066; DB 2; Length 3066;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3066; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAAACAGCTATGACCATGATTACGCCAAGCTCGAAATTAACCTCTCACTAAGGGAAACAA 60
Db 1 GGAAACAGCTATGACCATGATTACGCCAAGCTCGAAATTAACCTCTCACTAAGGGAAACAA 60
Qy 61 AAGCTGGAGCTCCACCGGGTGGCGGCTCTAGAACTAGTGGATCCCGGGCTGCAG 120
Db 61 AAGCTGGAGCTCCACCGGGTGGCGGCTCTAGAACTAGTGGATCCCGGGCTGCAG 120
Qy 121 GAAATTCGGCAGAGGCTCCACCGACAGCCAGGCACTGGGCGACGCACTGGAGACCCAG 180
Db 121 GAAATTCGGCAGAGGCTCCACCGACAGCCAGGCACTGGGCGACGCACTGGAGACCCAG 180
Qy 181 GACCTGTGAGGAGCAGCTCCGGGTGACAGGGGAGCTGAAGATCTCCACAGGGGC 240
Db 181 GACCTGTGAGGAGCAGCTCCGGGTGACAGGGGAGCTGAAGATCTCCACAGGGGC 240
Qy 241 TCACGAGGACAAATGGGTAAACCAATGAGTCTCCCAAGAGTTGAAGCAAGAGAAT 300
Db 241 TCACGAGGACAAATGGGTAAACCAATGAGTCTCCCAAGAGTTGAAGCAAGAGAAT 300
Qy 301 GAAACAGAGCAGAGACTTACAGGACAAACCGCTCTGCTCTGAACGGGGTTCAGTGGTG 360
Db 301 GAAACAGAGCAGAGACTTACAGGACAAACCGCTCTGCTCTGAACGGGGTTCAGTGGTG 360
Qy 361 GTGTGACCCACACAGTTCAGACTTTAGAGAAAGTCCGACTTGGGAATTAAGTGTCAAGACG 420
Db 361 GTGTGACCCACACAGTTCAGACTTTAGAGAAAGTCCGACTTGGGAATTAAGTGTCAAGACG 420
Qy 421 GATATGTGGCCACTTCTTCCCGGAGACACGAGATAGTGTCTTGGGATGCCAAC 480
Db 421 GATATGTGGCCACTTCTTCCCGGAGACACGAGATAGTGTCTTGGGATGCCAAC 480
Qy 481 GGAAAGAAATCTTGGGAAGAGGCCAAACCCGAGGACCAAGCTGCTAAATCTCGTTTTTC 540
Db 481 GGAAAGAAATCTTGGGAAGAGGCCAAACCCGAGGACCAAGCTGCTAAATCTCGTTTTTC 540
Qy 541 TTGATGCTCTCTCGGCTGTACAGGAGCTACCGGAGACCAAGCGCGAGATTCATCCCTT 600
Db 541 TTGATGCTCTCTCGGCTGTACAGGAGCTACCGGAGACCAAGCGCGAGATTCATCCCTT 600
Qy 601 GGATCAGTGAAGCTTGAATGTCAGCTCCCAATAGCTCCAGGACCAAGACCCCAAGTGAG 660
Db 601 GGATCAGTGAAGCTTGAATGTCAGCTCCCAATAGCTCCAGGACCAAGACCCCAAGTGAG 660
Qy 661 AGCTGGACACTTCCGGTGGCAGCTGGACCGGGCGAGGACACAGATAAAACCCAGGGCAC 720
Db 661 AGCTGGACACTTCCGGTGGCAGCTGGACCGGGCGAGGACACAGATAAAACCCAGGGCAC 720
Qy 721 GCCCGGCCCAAGACAAAGGCTCTCTGCGCGCAGGGATCCACGCTCTCCACCTGAG 780
Db 721 GCCCGGCCCAAGACAAAGGCTCTCTGCGCGCAGGGATCCACGCTCTCCACCTGAG 780
Qy 781 ACAGGGGAGCAGGAGGAGAGCTCCCTCCAGCCCAAGGAGCTCCAGCTTTTGGACAAA 840
Db 781 ACAGGGGAGCAGGAGGAGAGAGCTCCCTCCAGCCCAAGGAGCTCCAGCTTTTGGACAAA 840

Qy 841 TTCTTCAAGCTGACAAAGGACAGGAAAGGTCAGGTGACAGCCAAACAGGAAGCCAAG 900
Db 841 TTCTTCAAGCTGACAAAGGACAGGAAAGGTCAGGTGACAGCCAAACAGGAAGCCAAG 900
Qy 901 AGGCGAGAGCATCAAGACAAAGGTGGATGAGGTTCTGGCTTATCAGGCGAGTCCGATGAT 960
Db 901 AGGCGAGAGCATCAAGACAAAGGTGGATGAGGTTCTGGCTTATCAGGCGAGTCCGATGAT 960
Qy 961 GTCCCTCAGGGAAGGACATAGTTGACGCAAGGAAAGAAAGGACAAAGAACTTGGAACT 1020
Db 961 GTCCCTCAGGGAAGGACATAGTTGACGCGCAAGGAAAGAAAGGACAAAGAACTTGGAACT 1020
Qy 1021 GCGGATTGCTCTGTCTCTGGGACCCAGAAAGGACTGGAGACTGCAAAAGGACGATCCAG 1080
Db 1021 GCGGATTGCTCTGTCTCTGGGACCCAGAAAGGACTGGAGACTGCAAAAGGACGATCCAG 1080
Qy 1081 GCAGCAGCTATAGCAGAGAAATAATATCCATCATGAGTTCTTTAAACCTCTGGTTCA 1140
Db 1081 GCAGCAGCTATAGCAGAGAAATAATATTCATCATGAGTTCTTTAAACCTCTGGTTCA 1140
Qy 1141 CCTAAACAAAGCTGAAAACAAAAGGACCCAGAAAGGACACGGGTGCTGAAAAGTCAACCACC 1200
Db 1141 CCTAAACAAAGCTGAAAACAAAAGGACCCAGAAAGGACACGGGTGCTGAAAAGTCAACCACC 1200
Qy 1201 ACTTCAGCTGACCTTAAGTCAGACAAAGCCAACTTTTACATCCAGGAGACCCAAAGGGCT 1260
Db 1201 ACTTCAGCTGACCTTAAGTCAGACAAAGCCAACTTTTACATCCAGGAGACCCAAAGGGCT 1260
Qy 1261 GGCNAGAAATTCNAAAGGATGCNACCCATCGGGGCACACAGTCCGTGACAAACCCCTGAA 1320
Db 1261 GGCNAGAAATTCNAAAGGATGCNACCCATCGGGGCACACAGTCCGTGACAAACCCCTGAA 1320
Qy 1321 CCTCGAAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
Db 1321 CCTCGAAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
Qy 1381 TTTTGGAAAAAGTCAGTTAAAGAGGACTCAGTCCCAACAGTGCAGAGGAAATGTGGTG 1440
Db 1381 TTTTGGAAAAAGTCAGTTAAAGAGGACTCAGTCCCAACAGTGCAGAGGAAATGTGGTG 1440
Qy 1441 TGTGAGTCACAGTAGAGATTATAAGTCCNAGAAAGTAGNATCAGCTTTACAAACAGTG 1500
Db 1441 TGTGAGTCACAGTAGAGATTATAAGTCCNAGAAAGTAGNATCAGCTTTACAAACAGTG 1500
Qy 1501 GACCTCAACGAAGAGATGCTGCACTTGAACCCACAGAAAGGAACTCAAAAGGAGAA 1560
Db 1501 GACCTCAACGAAGAGATGCTGCACTTGAACCCACAGAAAGGAACTCAAAAGGAGAA 1560
Qy 1561 AGCAAAACAAAGAACTCTCTGATGGCGTTTCTCAGACAAATGTGAGTGAAGGGATGGA 1620
Db 1561 AGCAAAACAAAGAACTCTCTGATGGCGTTTCTCAGACAAATGTGAGTGAAGGGATGGA 1620
Qy 1621 GGGATCACCCACTCAGAAAGAAATAAATGGGAAAGACTCCAGTCCCAAACTCAGACTCC 1680
Db 1621 GGGATCACCCACTCAGAAAGAAATAAATGGGAAAGACTCCAGTCCCAAACTCAGACTCC 1680
Qy 1681 ACAGAAAAGACTATCACAACCGCCAGAGCTGAAACCAACAGGAGGACCCACAGAGGGTAA 1740
Db 1681 ACAGAAAAGACTATCACAACCGCCAGAGCTGAAACCAACAGGAGGACCCACAGAGGGTAA 1740
Qy 1741 GAGGGCTCTCTGAAAGGACAAAGATCAGAGCCGAGATGAACAAGAGAGAAAGCAAG 1800
Db 1741 GAGGGCTCTCTGAAAGGACAAAGATCAGAGCCGAGATGAACAAGAGAGAAAGCAAG 1800
Qy 1801 CAGCAAGCCCAAGAAACAGCCAGTGCACAGAGAGGACCGGTGGACCACTCAGCTG 1860
Db 1801 CAGCAAGCCCAAGAAACAGCCAGTGCACAGAGAGGACCGGTGGACCACTCAGCTG 1860
Qy 1861 CAGAAATGGGACAAAGCTCCAAAAGAGACTTGAAGAGGGCAGAGTCCCTTTGGGGCTTC 1920
Db 1861 CAGAAATGGGACAAAGCTCCAAAAGAGACTTGAAGAGGGCAGAGTCCCTTTGGGGCTTC 1920
Qy 1921 TTTAAAGGCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAACAGACCCAGTATCC 1980

```
Db 1921 TTTAAAGGCTGGGACCAAGCGGATGTTGATGCTCAAGTGCAAAAGAGCCAGTATCC 1980
Qy 1981 ATCGGACAGTTGGCAAAACCCAAAGTAAACAAATCAGCAGCGTTCCACAGGTTCTCTG 2040
Db 1981 ATCGGACAGTTGGCAAAACCCAAAGTAAACAAATCAGCAGCGTTCCACAGGTTCTCTG 2040
Qy 2041 CCACCAAGATGTCTCTCTTACTCCATCTCTCTCCCAACACGCTCCATGTATATTC 2100
Db 2041 CCACCAAGATGTCTCTCTTACTCCATCTCTCTCCCAACACGCTCCATGTATATTC 2100
Qy 2101 TTCTGATGCCAGCAAAATCTGCTAGAAATTAAGCCGAGCTGTTGTATATGA 2160
Db 2101 TTCTGATGCCAGCAAAATCTGCTAGAAATTAAGCCGAGCTGTTGTATATGA 2160
Qy 2161 GGTGTATTTATTAAGTCTCTGTCAGTCTCTTTCTGCAAAATAACAGTAAAGATGTTTA 2220
Db 2161 GGTGTATTTATTAAGTCTCTGTCAGTCTCTTTCTGCAAAATAACAGTAAAGATGTTTA 2220
Qy 2221 GCAGTCACTAGTTGGGTCAGAGAGTCGATGATCACAAGCAGGAAGGGAGGAATA 2280
Db 2221 GCAGTCACTAGTTGGGTCAGAGAGTCGATGATCACAAGCAGGAAGGGAGGAATA 2280
Qy 2281 GAGGAATGTGTTGGGTTAAGTATGATAAATGGCAGTGTGGCCGGGCTGTGCTCTC 2340
Db 2281 GAGGAATGTGTTGGGTTAAGTATGATAAATGGCAGTGTGGCCGGGCTGTGCTCTC 2340
Qy 2341 GCCTGTAAATCTCAGCAGTTTGGGAGCCGAGGAGTGATCACCTGAGGTCAGGAGTTC 2400
Db 2341 GCCTGTAAATCTCAGCAGTTTGGGAGCCGAGGAGTGATCACCTGAGGTCAGGAGTTC 2400
Qy 2401 AAGACTAGCTGGCCCAACATCATGAACCCCGTCTCTACTAAATAACAAAAATTAGCCA 2460
Db 2401 AAGACTAGCTGGCCCAACATCATGAACCCCGTCTCTACTAAATAACAAAAATTAGCCA 2460
Qy 2461 GGCATGTGGCAGACACACCTGTAGTCCAGTACTCGGAGCCCAACGACGAGAACCGCT 2520
Db 2461 GGCATGTGGCAGACACACCTGTAGTCCAGTACTCGGAGCCCAACGACGAGAACCGCT 2520
Qy 2521 TGTATCCAGGAGGTGGAGGTTGCAAGTGGAGCCGAAAGTTGCACCATTTGCACCTCCG 2580
Db 2521 TGTATCCAGGAGGTGGAGGTTGCAAGTGGAGCCGAAAGTTGCACCATTTGCACCTCCG 2580
Qy 2581 CGACAGACAGATTTCTATCAAAAAAAGAGCAGTGGCAAGTATAGAGAGAAA 2640
Db 2581 CGACAGACAGATTTCTATCAAAAAAAGAGCAGTGGCAAGTATAGAGAGAAA 2640
Qy 2641 TGCTGCTAGAAGGAATTAAGCGTTGCTAGTAAACGGTCTCATCTCTTAAGCTTGAAGA 2700
Db 2641 TGCTGCTAGAAGGAATTAAGCGTTGCTAGTAAACGGTCTCATCTCTTAAGCTTGAAGA 2700
Qy 2701 GGGAGACGAAATCCATTTGTTTAAATTCACATCTCAAGGAGGGAGAACCCGGGCTGTGT 2760
Db 2701 GGGAGACGAAATCCATTTGTTTAAATTCACATCTCAAGGAGGGAGAACCCGGGCTGTGT 2760
Qy 2761 TGGGTGTTGCCAATTTCTAGACGGAATGTGTGGGTATAGAAAAAGGAATGAATAAG 2820
Db 2761 TGGGTGTTGCCAATTTCTAGACGGAATGTGTGGGTATAGAAAAAGGAATGAATAAG 2820
Qy 2821 CGTTGTTTTTCAATAGGCTCTGTAGTTATTGATGAGAGGGAAGAGATTGACTGGGG 2880
Db 2821 CGTTGTTTTTCAATAGGCTCTGTAGTTATTGATGAGAGGGAAGAGATTGACTGGGG 2880
Qy 2881 AGGGCTTAAATGATTTGGGAAAAACAATTTGCTTTTGGGCTCAGTGACAAACGGCAAGAT 2940
Db 2881 AGGGCTTAAATGATTTGGGAAAAACAATTTGCTTTTGGGCTCAGTGACAAACGGCAAGAT 2940
Qy 2941 TACAATTTAAAAAATAAAAAAATAAAAACTCGAGACTAGTTCTCTCTCTCTCTGTCGCGA 3000
Db 2941 TACAATTTAAAAAATAAAAAAATAAAAACTCGAGACTAGTTCTCTCTCTCTCTGTCGCGA 3000
Qy 3001 ATTCGATATCAAGCTTATCGATACCGTGCAGCTCGAGGGGGGGCCGGTATCCCAATTCGC 3060
Db 3001 ATTCGATATCAAGCTTATCGATACCGTGCAGCTCGAGGGGGGGCCGGTATCCCAATTCGC 3060
```

```
Db 3001 ATTCGATATCAAGCTTATCGATACCGTGCAGCTCGAGGGGGGGCCGGTATCCCAATTCGC 3060
Qy 3061 CCTATA 3066
Db 3061 CCTATA 3066

RESULT 2
ADQ23347
ID ADQ23347 standard; DNA; 3418 BP.
XX
AC ADQ23347;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 6167.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
ds.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
XX WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
of a gene in a first soft tissue sample and a normal soft tissue sample
and comparing the gene expression, also useful in treating soft tissue
sarcoma.
XX
PS Example 2; SEQ ID NO 6167; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
which comprises obtaining a first soft tissue sample from an individual
and a normal soft tissue sample from the same or different individual,
determining the expression of a gene in both samples and comparing the
expression of the gene in both soft tissue samples, where a higher level
of protein expression in the first soft tissue sample indicates the
presence of soft tissue sarcoma. The method of the invention has
cytostatic applications and may be useful for detecting soft tissue
sarcoma, possibly via gene therapy or vaccine production. The nucleic
acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
DNA of the invention. The current sequence is not shown within the
specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 3418 BP; 1029 A; 763 C; 842 G; 669 T; 0 U; 115 Other;

Query Match 92.3%; Score 2828.8; DB 12; Length 3418;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2833; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 127 GGACAGAGGCTCCACCGACAGCCAGGCACTGGGAGCAGCAGCTGGAGACCCAGGACCT 186
Db 40 GGAGCCAGGCTCCACCGACAGCCAGGCACTGGGAGCAGCAGCTGGAGACCCAGGACCT 99

Qy 187 GTGAGGAGCAGCTCCGGGTGACACGAGGGGACATGAAGATATCTCCACAGGGGCTCAGCA 246
Db 100 GTGAGGAGCAGCTCCGGGTGACACGAGGGGACATGAAGATATCTCCACAGGGGCTCAGCA 159

Qy 247 GGACCAATGGGTATACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAAATGAACCA 306
Db 247 GGACCAATGGGTATACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAAATGAACCA 306
```

Db 160 GGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAAATGAACCA 219
Qy 307 GAACAGAGACTTACCAAGGACCAACCGTCTGCTCTGAACGGGTTCAGTGGTGTGTCG 366
Db 220 GAACAGAGACTTACCAAGGACCAACCGTCTGCTCTGAACGGGTTCAGTGGTGTGTCG 279
Qy 367 ACCACACAGTTCAGCACTTAGAGAAAGTGCAGCTTGGGAATAAGTGTCAAGACGGATAAT 426
Db 280 ACCACACAGTTCAGCACTTAGAGAAAGTGCAGCTTGGGAATAAGTGTCAAGACGGATAAT 339
Qy 427 GTGGCCACTTCTCCCGGACAGCAACCGAGATAGTGTGTGGGATGCCAACGGAAAG 486
Db 340 GTGGCCACTTCTCCCGGACAGCAACCGAGATAGTGTGTGGGATGCCAACGGAAAG 399
Qy 487 AATCTTGGGAAGAGGCCAAACCCGAGGACCAAGCTGTCTAAATCTCGTTTTTTCTTGATG 546
Db 400 AATCTTGGGAAGAGGCCAAACCCGAGGACCAAGCTGTCTAAATCTCGTTTTTTCTTGATG 459
Qy 547 CTCTCTCGGCTGTACCAAGGACGTACCGGAGACCAAGCCGAGATTATCCCTTTGGATCA 606
Db 460 CTCTCTCGGCTGTACCAAGGACGTACCGGAGACCAAGCCGAGATTATCCCTTTGGATCA 519
Qy 607 GTGAAGCTTGATGTGACGTCCAAATAAGCTCCAGCGCAACAAAGACCCAAAGTGAGAGCTGG 666
Db 520 GTGAAGCTTGATGTGACGTCCAAATAAGCTCCAGCGCAACAAAGACCCAAAGTGAGAGCTGG 579
Qy 667 ACATCTCCGCTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCAGGSCACGCCCG 726
Db 580 ACATCTCCGCTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCAGGSCACGCCCG 639
Qy 727 GCCCAAGAACAAAGTCTCTCTGCGGCCAGGATCCACGCTTCTCCACCTTGACACAGG 786
Db 640 GCCCAAGAACAAAGTCTCTCTGCGGCCAGGATCCACGCTTCTCCACCTTGACACAGG 699
Qy 787 GGAGCAGAGAGAGTCCCTCCAGCCCAAGAGTCCAGCTTTTGTGACAAATTTCTTC 846
Db 700 GGAGCAGAGAGAGTCCCTCCAGCCCAAGAGTCCAGCTTTTGTGACAAATTTCTTC 759
Qy 847 AAGCTGGACAGGACAGGAAAGGTGCCAGTGCACAGCCAAAGGAGGCCAAAGGGCA 906
Db 760 AAGCTGGACAGGACAGGAAAGGTGCCAGTGCACAGCCAAAGGAGGCCAAAGGGCA 819
Qy 907 GAGCATCAAGACAAGTGTGATGAGTTCCTGGCTTATCAGGCGAGTCCGATGATGTCCCT 966
Db 820 GAGCATCAAGACAAGTGTGATGAGTTCCTGGCTTATCAGGCGAGTCCGATGATGTCCCT 879
Qy 967 GCAGGGAAGACATAGTTGACGGCAAGGAAAGAGAGCAAGAACTTGGAACTGCCGAT 1026
Db 880 GCAGGGAAGACATAGTTGACGGCAAGGAAAGAGAGCAAGAACTTGGAACTGCCGAT 939
Qy 1027 TGCTCTGCTCCCTGGGACCCAGAGGACTGGAGACTGCAAGGACGATTCCAGGCA 1086
Db 940 TGCTCTGCTCCCTGGGACCCAGAGGACTGGAGACTGCAAGGACGATTCCAGGCA 999
Qy 1087 GCTATAGCAGAGAATAAATTCATCATGAGTTTCTTTAAACTCTGGTTTCACTTAAC 1146
Db 1000 GCTATAGCAGAGAATAAATTCATCATGAGTTTCTTTAAACTCTGGTTTCACTTAAC 1059
Qy 1147 AAAGCTGAAACAAAAGGACCCAGAGACAGCGGTGTGAAAAGTCAACCACTTCA 1206
Db 1060 AAAGCTGAAACAAAAGGACCCAGAGACAGCGGTGTGAAAAGTCAACCACTTCA 1119
Qy 1207 GCTGACCTTAAGTCAGACAAAGCCAACTTTATCTCCAGGAGACCAAGGGGCTGGCAAG 1266
Db 1120 GCTGACCTTAAGTCAGACAAAGCCAACTTTATCTCCAGGAGACCAAGGGGCTGGCAAG 1179
Qy 1267 AATTCCAAAGGATGCAACCCCTCGGGGACACACAGTCCGTGCAACCCCTGAACTTCG 1326
Db 1180 AATTCCAAAGGATGCAACCCCTCGGGGACACACAGTCCGTGCAACCCCTGAACTTCG 1239
Qy 1327 AAGGAGGACCAAGGAGAAATCAGGACCCACTCTCTGGGCAAACTGTGTTGG 1386
Db 1240 AAGGAGGACCAAGGAGAAATCAGGACCCACTCTCTGGGCAAACTGTGTTGG 1299

Qy 1387 AAAAAGTCAGTTAAAGAGGACTCAGTCCCAAGGTCGGAGGAGAAATGTGGTGTGAG 1446
Db 1300 AAAAAGTCAGTTAAAGAGGACTCAGTCCCAAGGTCGGAGGAGAAATGTGGTGTGAG 1359
Qy 1447 TCACCAAGTAGAGATTATAAGTCCAAGAAAGTGAATCAGCCTTACAAACAGTGGACCTC 1506
Db 1360 TCACCAAGTAGAGATTATAAGTCCAAGAAAGTGAATCAGCCTTACAAACAGTGGACCTC 1419
Qy 1507 AACGAGGAGATCTGCACCTGACCCACAGAACGCAAACTCAAAACAGAGAAAGCAAA 1566
Db 1420 AACGAGGAGATCTGCACCTGACCCACAGAACGCAAACTCAAAACAGAGAAAGCAAA 1479
Qy 1567 CCAAGAAACCTCTCTGATGGGCTTTCTCAGACAAATGTCAAGAAAGGGATGGAGGATC 1626
Db 1480 CCAAGAAACCTCTCTGATGGGCTTTCTCAGACAAATGTCAAGAAAGGGATGGAGGATC 1539
Qy 1627 ACCCACTCAGAGAAATAAATGGGAAAGATCTCAGCTGCCAAACATCAGACTCCACAGAA 1686
Db 1540 ACCCACTCAGAGAAATAAATGGGAAAGATCTCAGCTGCCAAACATCAGACTCCACAGAA 1599
Qy 1687 AAGACTATCACACCGCCAGAGCCTGAACCAACAGGAGCACCAAGAGGGTAAAGAGGC 1746
Db 1600 AAGACTATCACACCGCCAGAGCCTGAACCAACAGGAGCACCAAGAGGGTAAAGAGGC 1659
Qy 1747 TCCTCGAAGGACAAGAAAGTCAGCAGCCGAGATGAACCAAGCAGAGAGAGCAACAAGCAGGAA 1806
Db 1660 TCCTCGAAGGACAAGAAAGTCAGCAGCCGAGATGAACCAAGCAGAGAGAGCAACAAGCAGGAA 1719
Qy 1807 GCCAAAGAAACAGCCCAAGTGCACAGAGAGCCACGGTGGACAGAACTCAGTCAGAAAT 1866
Db 1720 GCCAAAGAAACAGCCCAAGTGCACAGAGAGCCACGGTGGACAGAACTCAGTCAGAAAT 1779
Qy 1867 GGGGACAAGCTCCAAAGAGAGACCTGGAAGCGGACAGTCCCTTGGGGCTCTTTTAA 1926
Db 1780 GGGGACAAGCTCCAAAGAGAGACCTGGAAGCGGCNAGGTCCTTGGGGCTCTTTTAA 1839
Qy 1927 GGCTCGGACCAAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGA 1986
Db 1840 GGCTCGGACCAAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGA 1899
Qy 1987 CCAGTTGGCAAAACCCAAAGTAAACAAATCAGACAGCGTTCCCAAGAGTTCTCTGCCACCA 2046
Db 1900 CCAGTTGGCAAAACCCAAAGTAAACAAATCAGACAGCGTTCCCAAGAGTTCTCTGCCACCA 1959
Qy 2047 AGATGTCTCTCTTACTCCATCTCTCCCAACACGCTCCATGTATATATTCTTCTGA 2106
Db 1960 AGATGTCTCTCTTACTCCATCTCTCCCAACACGCTCCATGTATATATTCTTCTGA 2019
Qy 2107 TGGCCAGCAAAATGAAATTTGCTCCTAGAAATTTAAGCCCGAGCTGTTGTATATTGAGGTGTA 2166
Db 2020 TGGCCAGCAAAATGAAATTTGCTCCTAGAAATTTAAGCCCGAGCTGTTGTATATTGAGGTGTA 2079
Qy 2167 TTATTTACGTCTCTGGTCCAGTCTTTTCTGCAAAATAACAGTAAAGATGGTTTACAGGT 2226
Db 2080 TTATTTACGTCTCTGGTCCAGTCTTTTCTGCAAAATAACAGTAAAGATGGTTTACAGGT 2139
Qy 2227 CACCTAGTTGGTCAGAAAGAGTCGATGATCACCAGCAGGAAAGGGAGGAATAGAGAA 2286
Db 2140 CACCTAGTTGGTCAGAAAGAGTCGATGATCACCAGCAGGAAAGGGAGGAATAGAGAA 2199
Qy 2287 TGTCTTCCGGTTAAGTGAATAAATGGCAGTGGTGGCCGGCGTGGTGGCTCTCGCCTGT 2346
Db 2200 TGTCTTCCGGTTAAGTGAATAAATGGCAGTGGTGGCCGGCGTGGTGGCTCTCGCCTGT 2259
Qy 2347 AATCTCAGACTTTGGGAGGCCGAGGCGAGTGGATCACCCTGAGTTCAGGAGTCAAGACT 2406
Db 2260 AATCTCAGACTTTGGGAGGCCGAGGCGAGTGGATCACCCTGAGTTCAGGAGTCAAGACT 2319
Qy 2407 AGCCTGGCACAACATCATGAAACCCGCTCTCTACTTAAATAACAAAAATTTAGCCAGGCAATG 2466
Db 2320 AGCCTGGCACAACATCATGAAACCCGCTCTCTACTTAAATAACAAAAATTTAGCCAGGCAATG 2379

QY 2467 GTGGCACACCTGTAGTCCAGCTACTCGGAGGCCAACGACGAGAACCGCTTGTACC 2526
DB 2380 GTGGCACACCTGTAGTCCAGCTACTCGGAGGCCAACGACGAGAACCGCTTGTACC 2439
QY 2527 CAGGAGGTGGAGTTGCACTGAGCCGAGAAAGTTGACCATTCACCTCCACCTGGGGACAG 2586
DB 2440 CAGGAGGTGGAGTTGCACTGAGCCGAGAAAGTTGACCATTCACCTCCACCTGGGGACAG 2499
QY 2587 ACCAGATTCATCAAAAAAAGGCGAGTGCACAGTAAGTTATAGAGAGAAATGCTGC 2646
DB 2500 AGCAAGATTCATCAAAAAAAGGCGAGTGCACAGTAAGTTATAGAGAGAAATGCTGC 2559
QY 2647 TAGAGGAATTAAGCGTTGTAGTAAACCGGTCTCATCTCTAAAGCTTGAAGAGGGAGA 2706
DB 2560 TAGAGGAATTAAGCGTTGTAGTAAACCGGTCTCATCTCTAAAGCTTGAAGAGGGAGA 2619
QY 2707 CGAAATCCATTTGTTAAATTCACATCTCAAGAGGGAGAACCCGGGCTGTGTTGGGTG 2766
DB 2620 CGAAATCCATTTGTTAAATTCACATCTCAAGAGGGAGAACCCGGGCTGTGTTGGGTG 2679
QY 2767 GTTGCAATTTCTAGAACGGAATGTGGGGTATAGAAAAAGGAATGAATAGCGTTGT 2826
DB 2680 GTTGCAATTTCTAGAACGGAATGTGGGGTATAGAAAAAGGAATGAATAGCGTTGT 2739
QY 2827 TTTTCAATAGGCTCTTGTAGTTATTGATGAGAGGGAAGAAAGATTGACTGGGAGGCT 2886
DB 2740 TTTTCAATAGGCTCTTGTAGTTATTGATGAGAGGGAAGAAAGATTGACTGGGAGGCT 2799
QY 2887 TAAATGATTTGGGAAAAAATTTGCTTTTGGGCTCAGTGACAAACGGCAAGATTACAAAC 2946
DB 2800 TAAATGATTTGGGAAAAAATTTGCTTTTGGGCTCAGTGACAAACGGCAAGATTACAAAC 2859
QY 2947 TTAATAAAAAA 2967
DB 2860 TTAATAAAAAAATAAATAA 2880

RESULT 3
AAS81115
ID AAS81115 standard; cDNA; 2813 BP.
XX
AC AAS81115;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #16919.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ASB16928.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
XX biodiversity.

PS Claim 1; SEQ ID NO 16919; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 2813 BP; 868 A; 671 C; 758 G; 516 T; 0 U; 0 Other;
Query Match 91.7%; Score 2813; DB 5; Length 2813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 136 CTCACACGACAGCAGGCACTGGGCGAGCACCCACTGGAGACCAGGACCCCTGTGCAGGAG 195
DB 1 CTCACACGACAGCAGGCACTGGGCGAGCACCCACTGGAGACCAGGACCCCTGTGCAGGAG 60
QY 196 CAGCTCCCGGTGCACACGAGGGGACTGAAAGATACTCCACAGGGGCTCAGCAGGAGCAATG 255
DB 61 CAGCTCCCGGTGCACACGAGGGGACTGAAAGATACTCCACAGGGGCTCAGCAGGAGCAATG 120
QY 256 GGTAAACCAATGAGTGTTCCTCAAGAGTTGAAGACCAAGAGATGAACCCAGACAGAG 315
DB 121 GGTAAACCAATGAGTGTTCCTCAAGAGTTGAAGACCAAGAGATGAACCCAGACAGAG 180
QY 316 ACTTACAGGACACAGCGGTCTGCTGAAACGGGGTTCAGTGGTGTCCGACCCACACACA 375
DB 181 ACTTACAGGACACAGCGGTCTGCTGAAACGGGGTTCAGTGGTGTCCGACCCACACACA 240
QY 376 GTTCAGCACTTAGAGGAAGTCGACTTGGGAATAGTGTCAAGACGGATTAATGTGCCACT 435
DB 241 GTTCAGCACTTAGAGGAAGTCGACTTGGGAATAGTGTCAAGACGGATTAATGTGCCACT 300
QY 436 TCTTCCCCCGAGACAAACGGAGATAAGTGTCTGTGGGATGCCAACCGGAAAGAAATCTTGGG 495
DB 301 TCTTCCCCCGAGACAAACGGAGATAAGTGTCTGTGGGATGCCAACCGGAAAGAAATCTTGGG 360
QY 496 AAGAGGCCAAACCCGAGGACCCAGCTGCTAAATCTCGTTTTTCTTATGCTCTCTCGG 555
DB 361 AAGAGGCCAAACCCGAGGACCCAGCTGCTAAATCTCGTTTTTCTTATGCTCTCTCGG 420
QY 556 CCTGTACCGAGACGTACCGGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTT 615
DB 421 CCTGTACCGAGACGTACCGGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTT 480
QY 616 GATGTGAGTCCCAATAAAGCTCCAGCGAACCAAGACCAAGTGAGAGTGGACACTTCG 675
DB 481 GATGTGAGTCCCAATAAAGCTCCAGCGAACCAAGACCAAGTGAGAGTGGACACTTCG 540
QY 676 GTGCGAGTGCACCGGGCAGGACACAGATAAAACCCAGGCGACGCCGCCCAAGAC 735
DB 541 GTGCGAGTGCACCGGGCAGGACACAGATAAAACCCAGGCGACGCCGCCCAAGAC 600
QY 736 AAGTGTCTCTCTGCGGCCAGGAGTCCACAGCTTCTTCCCACTTGACAGAGGGGACAGGA 795
DB 601 AAGTGTCTCTCTGCGGCCAGGAGTCCACAGCTTCTTCCCACTTGACAGAGGGGACAGGA 660

Qy 796 GGAGAAGCTCCCTCCAAAGCCCAAGGACTCCAGCTTTTTTGTACAAATCTTCAAGCTGGAC 855
Db |||||
Qy 661 GGAGAGCTCCTCCAGGCCCAAGGACTCCAGCTTTTTTGTACAAATCTTCAAGCTGGAC 720
Db |||||
Qy 856 AAGGACAGGAAAAGGTGCCAGGTGACAGCCAAAGCAAGCAAGCAAGGAGGAGAGCATCAA 915
Db |||||
Qy 721 AAGGACAGGAAAAGGTGCCAGGTGACAGCCAAAGCAAGCAAGGAGGAGAGCATCAA 780
Db |||||
Qy 916 GACAAGTGGATGAGGTTCCTGGCTTATCAGGCGAGTCCGATGATGTCCCTGCAGGGAAG 975
Db |||||
Qy 781 GACAAGTGGATGAGGTTCCTGGCTTATCAGGCGAGTCCGATGATGTCCCTGCAGGGAAG 840
Db |||||
Qy 976 GACATAGTTGACGGCAAGGAAAAGAGGACAAAGAACTTTGGAACTTGCGGATTTGCTCTGTC 1035
Db |||||
Qy 841 GACATAGTTGACGGCAAGGAAAAGAGGACAAAGAACTTTGGAACTTGCGGATTTGCTCTGTC 900
Db |||||
Qy 1036 CCTGGGACCCAGAAAGACTGGAGACTGCAAGGACGATTCAGGAGCAGCAGCTATAGCA 1095
Db |||||
Qy 901 CCTGGGACCCAGAAAGACTGGAGACTGCAAGGACGATTCAGGAGCAGCAGCTATAGCA 960
Db |||||
Qy 1096 GAGATTAATAATCCATCAGTCTCTTTTAAACTCTGGTTTCACTTAAAGCAAGCTGAA 1155
Db |||||
Qy 961 GAGATTAATAATCCATCAGTCTCTTTTAAACTCTGGTTTCACTTAAAGCTGAA 1020
Db |||||
Qy 1156 AAAAAAGGACCCAGAAAGACTGGAGACTGCAAGGACGATTCAGGAGCAGCAGCTATAGCA 1215
Db |||||
Qy 1021 AAAAAAGGACCCAGAAAGACTGGAGACTGCAAGGACGATTCAGGAGCAGCAGCTATAGCA 1080
Db |||||
Qy 1216 AAGTCAGACAAAGCCAACTTTACATCCAGAGAGCCCAAGGGCTGGCAAGAAATCCAAA 1275
Db |||||
Qy 1081 AAGTCAGACAAAGCCAACTTTACATCCAGAGAGCCCAAGGGCTGGCAAGAAATCCAAA 1140
Db |||||
Qy 1276 GGATGCAACCATCGGGGACACACAGTCCCTGACACCCCTGACCTGCGAAGGAGGC 1335
Db |||||
Qy 1141 GGATGCAACCATCGGGGACACACAGTCCCTGACACCCCTGACCTGCGAAGGAGGC 1200
Db |||||
Qy 1336 ACCAAGGAGAAATCAGGACCCACCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1395
Db |||||
Qy 1201 ACCAAGGAGAAATCAGGACCCACCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1260
Db |||||
Qy 1396 GTTAAAGAGGACTCAGTCCCCACAGGTGCGAGGAGAAATGTTGTGATGATCAACAGTA 1455
Db |||||
Qy 1261 GTTAAAGAGGACTCAGTCCCCACAGGTGCGAGGAGAAATGTTGTGATGATCAACAGTA 1320
Db |||||
Qy 1456 GAGATTATAAGTCCAAGGAAGTAGAATCAGCCTTACAAACAGTGGACCTCAACGAGGA 1515
Db |||||
Qy 1321 GAGATTATAAGTCCAAGGAAGTAGAATCAGCCTTACAAACAGTGGACCTCAACGAGGA 1380
Db |||||
Qy 1516 GATGTCACCTGAAACCCACAGAGCGAAAACCTCAAAAGAGAAAGCAAAACCAAGAAC 1575
Db |||||
Qy 1381 GATGTCACCTGAAACCCACAGAGCGAAAACCTCAAAAGAGAAAGCAAAACCAAGAAC 1440
Db |||||
Qy 1576 TCTCTGATGGCGTTTCTCAGACAAATGTGAGTAAAGGGATGAGGGATCAACCCACTCA 1635
Db |||||
Qy 1441 TCTCTGATGGCGTTTCTCAGACAAATGTGAGTAAAGGGATGAGGGATCAACCCACTCA 1500
Db |||||
Qy 1636 GAAGAAATAATGGGAAAGACTCAGCTGCCAAGCATCAGACTCCAGAGAAAGACTATC 1695
Db |||||
Qy 1501 GAAGAAATAATGGGAAAGACTCAGCTGCCAAGCATCAGACTCCAGAGAAAGACTATC 1560
Db |||||
Qy 1696 ACACCGCAGAGCTGAAACCAACAGGAGCACCACAGAGGGGTAAAGAGGGCTCCTCGAAG 1755
Db |||||
Qy 1561 ACACCGCAGAGCTGAAACCAACAGGAGCACCACAGAGGGGTAAAGAGGGCTCCTCGAAG 1620
Db |||||
Qy 1756 GACAAGAAATCAGCAGCCGAGATGAACAGCAGAGAGCAACAGCAGGAAAGCCAAAGAA 1815
Db |||||
Qy 1621 GACAAGAAATCAGCAGCCGAGATGAACAGCAGAGAGCAACAGCAGGAAAGCCAAAGAA 1680
Db |||||
Qy 1816 CCAGCCAGTGCACAGAGCAGCCAGCTGGACAGCACTCACTGCAAGATGGGAGCAAG 1875
Db |||||
Qy 1681 CCAGCCAGTGCACAGAGCAGCCAGCTGGAGCAGCACTCACTGCAAGATGGGAGCAAG 1740
Db |||||

Qy 1876 CTCAAAAAGAGACTCTGAGAGCGGACAGTCCCTTGGGGCTTCTTTAAAGGCTGGGA 1935
Db |||||
Qy 1741 CTCAAAAAGAGACTCTGAGAGCGGACAGTCCCTTGGGGCTTCTTTAAAGGCTGGGA 1800
Db |||||
Qy 1936 CCAAAGCGGATGTTGGATGCTCAAGTCAAAACAGACCAGTATCCATCGGACCAAGTGGC 1995
Db |||||
Qy 1801 CCAAAGCGGATGTTGGATGCTCAAGTCAAAACAGACCAGTATCCATCGGACCAAGTGGC 1860
Db |||||
Qy 1996 AAACCCAAAGTAAACAAATCAGCAGGTTCCACCAGGTTCTCTGCCACCAAGATGTGTT 2055
Db |||||
Qy 1861 AAACCCAAAGTAAACAAATCAGCAGGTTCCACAGGTTCTCTGCCACCAAGATGTGTT 1920
Db |||||
Qy 2056 CTCCTTACTCCATCTCCTCCCAAAACACGCTCCATATATATATTTCTTCTGATGCCAGCA 2115
Db |||||
Qy 1921 CTCCTTACTCCATCTCCTCCCAAAACACGCTCCATATATATATTTCTTCTGATGCCAGCA 1980
Db |||||
Qy 2116 AATGAAATTTGCTCCTAGAAATTAAGCCCGAGCTGTTGTATATATTCAGGTGATATTTACG 2175
Db |||||
Qy 1981 AATGAAATTTGCTCCTAGAAATTAAGCCCGAGCTGTTGTATATTCAGGTGATATTTACG 2040
Db |||||
Qy 2176 TCTCTGTTCCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTTAGCAGGTCACTAGTT 2235
Db |||||
Qy 2041 TCTCTGTTCCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTTAGCAGGTCACTAGTT 2100
Db |||||
Qy 2236 GGGTCAGAAAGTCCATGATCAACCAAGCAGAAAGGAGGGAATAGAGGAATGTTCGG 2295
Db |||||
Qy 2101 GGGTCAGAAAGTCCATGATCAACCAAGCAGAAAGGAGGGAATAGAGGAATGTTCGG 2160
Db |||||
Qy 2296 GTTAAGTATGAAATGGCAGTGGTGGCGGGCGTGGTCTCGCTGCTTAATCTCAGC 2355
Db |||||
Qy 2161 GTTAAGTATGAAATGGCAGTGGTGGCGGGCGTGGTCTCGCTGCTTAATCTCAGC 2220
Db |||||
Qy 2356 ACTTTGGAGGCCGAGGCGAGTGGATCACTTGAGTCAAGGATTTCAAGACTAGCCTGGCC 2415
Db |||||
Qy 2221 ACTTTGGAGGCCGAGGCGAGTGGATCACTTGAGTCAAGGATTTCAAGACTAGCCTGGCC 2280
Db |||||
Qy 2416 AACATCATGAAACCCCGTCTCTACTTAAATAATACAAAAATAGCCAGGATGTTGGCACAC 2475
Db |||||
Qy 2281 AACATCATGAAACCCCGTCTCTACTTAAATAATACAAAAATAGCCAGGATGTTGGCACAC 2340
Db |||||
Qy 2476 ACCTGTAGTCCCAGTACTCGGAGCCCAAGCCACGAGAACCGCTTGTACCCAGGAGTG 2535
Db |||||
Qy 2341 ACCTGTAGTCCCAGTACTCGGAGCCCAAGCCACGAGAACCGCTTGTACCCAGGAGTG 2400
Db |||||
Qy 2536 GAGTTGAGTGAAGCCGAAAGTTGCACCATTTGCATCTCCACCTCGGCGCACAGCAAGATT 2595
Db |||||
Qy 2401 GAGTTGAGTGAAGCCGAAAGTTGCACCATTTGCATCTCCACCTCGGCGCACAGCAAGATT 2460
Db |||||
Qy 2596 CTATCAAAAAAAGGCGAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGTAGAGGAA 2655
Db |||||
Qy 2461 CTATCAAAAAAAGGCGAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGTAGAGGAA 2520
Db |||||
Qy 2656 TTAAAGCTTTGATTAAGCGGTGCTCATCTTAAGCTTGAAGAGGAGAGCAAGAAATCC 2715
Db |||||
Qy 2521 TTAAAGCTTTGATTAAGCGGTGCTCATCTTAAGCTTGAAGAGGAGAGCAAGAAATCC 2580
Db |||||
Qy 2716 ATTTGTTTAAATTCACATCTCAAGAGGAGAAACCCGGGCTGTGTGGGTGGTGGCCAAT 2775
Db |||||
Qy 2581 ATTTGTTTAAATTCACATCTCAAGAGGAGAAACCCGGGCTGTGTGGGTGGTGGCCAAT 2640
Db |||||
Qy 2776 TTCTTAGAACCGAAATGTGGGGTATAGAAAAAGAAATGAATAAGCGTTGTTTTCAAT 2835
Db |||||
Qy 2641 TTCTTAGAACCGAAATGTGGGGTATAGAAAAAGAAATGAATAAGCGTTGTTTTCAAT 2700
Db |||||
Qy 2836 AGGTCTTGTGATGATGAGGAGAAAGATTCACCTGGGAGGCTTTAAATGAT 2895
Db |||||
Qy 2701 AGGTCTTGTGATGATGAGGAGAAAGATTCACCTGGGAGGCTTTAAATGAT 2760
Db |||||
Qy 2896 TTGGAAAAAATTTGCTTTTGGGCTCAGTGACAAACGGCAAGAAAGATTACAAT 2948
Db |||||
Qy 2761 TTGGAAAAAATTTGCTTTTGGGCTCAGTGACAAACGGCAAGAAAGATTACAAT 2813
Db |||||


```

RESULT 4
AAD38092
ID AAD38092 standard; DNA; 2813 BP.
XX
AC AAD38092;
XX
DT 10-SEP-2002 (first entry)
XX
DE Human BCAS1 (breast cancer amplified sequence 1) DNA.
XX
KW Human; BCAS1; breast cancer amplified sequence 1; AIBCL1; inflammation;
KW amplified in breast cancer 1; NABCL1; novel amplified in breast cancer 1;
KW hyperproliferative disorder; breast; prostate; cancer; prophyllaxis;
KW infection; antisense therapy; cytostatic; antiinflammatory; tumour; gene;
KW ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 118..1872
FT FT /*tag= a
FT FT /product= "Human BCAS1 protein"
XX
PN WO200231136-A1.
XX
PD 18-APR-2002.
XX
PP 09-OCT-2001; 2001WO-US031484.
XX
PR 11-OCT-2000; 2000US-00689255.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Cowser LM, Freier SM;
XX
DR WPI; 2002-444179/47.
XX
DR P-PSDB; AAE23791.
XX
PT New antisense compounds targeted to a nucleic acid molecule encoding
PT BCAS1, useful for treating diseases or conditions associated with BCAS1,
PT such as hyperproliferative disease, particularly breast or prostate
PT cancer.
XX
PS Example 10; Page 92-95; 104pp; English.
XX
CC The invention relates to antisense compounds, compositions and methods
CC for modulating the expression of BCAS1 (breast cancer amplified sequence
CC 1, also known as AIBCL1 for amplified in breast cancer 1 and NABCL1 for
CC novel amplified in breast cancer 1). The antisense compounds of the
CC invention are useful for treating an animal having a disease or condition
CC associated with BCAS1, such as hyperproliferative disorders including
CC breast or prostate cancer. These compounds are also used as research
CC reagents and diagnostics; to distinguish between functions of various
CC members of a biological pathway; in the treatment of a disease or
CC disorder, which can be treated by modulating the expression of BCAS1; as
CC prophylaxis, e.g. to prevent or delay infection, inflammation or tumour
CC formation; and as probes or primers. These antisense compounds are used
CC in antisense therapy. The present sequence is human BCAS1 DNA
XX
SQ Sequence 2813 BP; 868 A; 671 C; 758 G; 516 T; 0 U; 0 Other;
XX
Query Match 91.7%; Score 2813; DB 6; Length 2813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 136 CTCACCCAGCCAGGCACTGGGAGCAGCCAGCTGGAGACCCAGGACCTGTGCAGGAG 195
DB 1 CTCACCCAGCAGCCAGGCACTGGGAGCAGCCAGCTGGAGACCCAGGACCTGTGCAGGAG 60
XX
QY 196 CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCCAAGGGGCTCAGCAGGAGCAATG 255
DB 61 CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCCAAGGGGCTCAGCAGGAGCAATG 120

```

[illegible]

| | | | | |
|----------|--|-----|---|------|
| Db | 2281 | AA | CATCATGAAACCCCGCTCTCTACTATAAAATACAAAATTTAGCAGCGCATGGTGCACAC | 2340 |
| Qy | 2476 | AC | CTGTAGTCCCGAGCTACTCGGAGGCCAACGACGACGAAACCGCTTTGTACCCAGGAGGTG | 2535 |
| Db | 2341 | AC | CTGTAGTCCCGAGCTACTCGGAGGCCAACGACGACGAAACCGCTTTGTACCCAGGAGGTG | 2400 |
| Qy | 2536 | GAG | TTGCACTGAGCCGAAGTTGGACCAATTTGCATCTCCACCTCTGGCGGACAGACGAAGATT | 2595 |
| Db | 2401 | GAG | TTGCACTGAGCCGAAGTTGGACCAATTTGCATCTCCACCTCTGGCGGACAGACGAAGATT | 2460 |
| Qy | 2596 | CT | ATCAAAAAAAGGCGAGTGCACAGTAAGTTATAGAAGAGAAATGCTCTGTAGAAGGAA | 2655 |
| Db | 2461 | CT | ATCAAAAAAAGGCGAGTGCACAGTAAGTTATAGAAGAGAAATGCTCTGTAGAAGGAA | 2520 |
| Qy | 2656 | TT | AAGCGTTGTAGTAAACCGGTGCTCATCTCTCTAAAGCTTCAAGAAAGGAGACGAAAAATCC | 2715 |
| Db | 2521 | TT | AAGCGTTGTAGTAAACCGGTGCTCATCTCTCTAAAGCTTCAAGAAAGGAGACGAAAAATCC | 2580 |
| Qy | 2716 | ATT | TGTTTAAATTCACATCTCAAGAGGGAGAAACCCGGGCTGTGTTGGGTGGTGGCCAAAT | 2775 |
| Db | 2581 | ATT | TGTTTAAATTCACATCTCAAGAGGGAGAAACCCGGGCTGTGTTGGGTGGTGGCCAAAT | 2640 |
| Qy | 2776 | TT | CTAGAACCGGAATGTGTGGGTATAGAAAAAGGAATGAATGACCTTGTTTTTCAAAT | 2835 |
| Db | 2641 | TT | CTAGAACCGGAATGTGTGGGTATAGAAAAAGGAATGAATGACCTTGTTTTTCAAAT | 2700 |
| Qy | 2836 | AG | GTCTCTTGTAAAGTTATTCATGAGAGGGAAAAAGATTGACTGGGGAGGGCTTAAAAATGAT | 2895 |
| Db | 2701 | AG | GTCTCTTGTAAAGTTATTCATGAGAGGGAAAAAGATTGACTGGGGAGGGCTTAAAAATGAT | 2760 |
| Qy | 2896 | TT | GGGAAAAACAATTTGCTTTTGGGCTCAGTGACACAAACGCAAGAAATTACAATT | 2948 |
| Db | 2761 | TT | GGGAAAAACAATTTGCTTTTGGGCTCAGTGACACAAACGCAAGAAATTACAATT | 2813 |
| RESULT 5 | | | | |
| ADJ75127 | | | | |
| ID | ADJ75127 standard; DNA; 2813 BP. | | | |
| AC | ADJ75127; | | | |
| XX | | | | |
| AC | | | | |
| XX | | | | |
| DT | 20-MAY-2004 (first entry) | | | |
| XX | | | | |
| DE | Marker gene SEQ ID NO:379. | | | |
| XX | | | | |
| XX | bronchial asthma; chronic obstructive pulmonary disease; | | | |
| KW | respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic; | | | |
| KW | gene therapy; marker gene; gene; ds. | | | |
| XX | | | | |
| OS | Homo sapiens. | | | |
| XX | | | | |
| PN | EPI394274-A2. | | | |
| XX | | | | |
| PD | 03-MAR-2004. | | | |
| PF | 04-AUG-2003; 2003EP-00254857. | | | |
| XX | | | | |
| PR | 06-AUG-2002; 2002JP-00229312. | | | |
| PR | 20-MAR-2003; 2003JP-00077212. | | | |
| XX | | | | |
| PA | (GENO-) GENOX RES INC. | | | |
| XX | | | | |
| PI | Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K; | | | |
| XX | | | | |
| DR | WPI; 2004-193155/19. | | | |
| XX | | | | |
| PT | Testing for bronchial asthma or chronic obstructive pulmonary disease by | | | |
| PT | comparing the expression level of a marker gene in a biological sample | | | |
| PT | from a subject with the expression level of the gene in a sample from a | | | |
| PT | healthy subject. | | | |
| XX | | | | |
| PS | Claim 1; SEQ ID NO 379; 241pp; English. | | | |
| XX | | | | |

Testing for bronchial asthma or chronic obstructive pulmonary disease by comparing the expression level of a marker gene in a biological sample from a subject with the expression level of the gene in a sample from a healthy subject.

Claim 1: SEO ID NO 379; 241pp; English.

The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (1) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (1) has respiratory and antiasthmatic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

Sequence 2813 BP; 868 A; 671 C; 758 G; 516 T; 0 U; 0 Other;

| Query Match | 91.7%; | Score 2813; | DB 12; | Length 2813; |
|-----------------------|--------------|---|------------|---------------------|
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 2813; | Conservative | 0; | Mismatches | 0; Indels 0; Gaps 0 |
| Qy | 136 | CTCCACCGACAGCCAGGCACTGGGCGAGCACGCACTGGAGACCCAGGACCCCTGTGCGAGGAG | 195 | |
| Db | 1 | CTCCACCGACAGCCAGGCACTGGGCGAGCACGCACTGGAGACCCAGGACCCCTGTGCGAGGAG | 60 | |
| Qy | 196 | CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCCAACAGGGGCTCAGCAGGACGAATG | 255 | |
| Db | 61 | CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCCAACAGGGGCTCAGCAGGACGAATG | 120 | |
| Qy | 256 | GGTAAACCAATGAGTGTTCCTCCCAAGAGGTTGAAGACCAAGAGATGAACCAAGACGAG | 315 | |
| Db | 121 | GGTAAACCAATGAGTGTTCCTCCCAAGAGGTTGAAGACCAAGAGATGAACCAAGACGAG | 180 | |
| Qy | 316 | ACTTTACAGGACAAACGCGTCTCTCTGAACGGGGTTCCAGTGGTGGTGTCGACCCACACA | 375 | |
| Db | 181 | ACTTACACAGACAAACGCGTCTCTCTGAACGGGGTTCCAGTGGTGGTGTCGACCCACACA | 240 | |
| Qy | 376 | GTTTCAGCACTTAGAGGAAGTGCATTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT | 435 | |
| Db | 241 | GTTTCAGCACTTAGAGGAAGTGCATTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT | 300 | |
| Qy | 436 | TCCTTCCCGCGAGCAACCGAGATAGTCTCTTGGGATGCCACCGAAGAGATCTTGGG | 495 | |
| Db | 301 | TCCTTCCCGCGAGCAACCGAGATAGTCTCTTGGGATGCCACCGAAGAGATCTTGGG | 360 | |
| Qy | 496 | AAAGAGGCCAAACCCGAGGACCAAGCTGCTAAATCTCGTTTTTCTTGATGCTCTCTCGG | 555 | |
| Db | 361 | AAAGAGGCCAAACCCGAGGACCAAGCTGCTAAATCTCGTTTTTCTTGATGCTCTCTCGG | 420 | |
| Qy | 556 | CCTGTATCAAGGACGTACCGGAGACCAAGCGCGAGATTCATCCCTTGGATCAGTGAAGCTT | 615 | |
| Db | 421 | CCTGTATCAAGGACGTACCGGAGACCAAGCGCGAGATTCATCCCTTGGATCAGTGAAGCTT | 480 | |
| Qy | 616 | GATGTCAAGTCCAAATAAGTCTCCACGGACCAAGACCCCAAGTGAAGCTCGACACTTCGG | 675 | |
| Db | 481 | GATGTCAAGTCCAAATAAGTCTCCACGGACCAAGACCCCAAGTGAAGCTCGACACTTCGG | 540 | |
| Qy | 676 | GTGGCAGCTGGGACCGGGGCGAGGACACAGATAAAACCCCAAGGACCGCCCGGCCCAAGAC | 735 | |

Db 1621 GACAGAAGTCAGCAGCCGAGATGAACAAGCAGAGAGCAACAAGCAGGAGCCAAAGAA 1680
Qy 1816 CCAGCCAGTGCACAGAGCAGCCACGCTGGACAGCAACTCAGTCAGAAATGGGGACAAG 1875
Db 1681 CCAGCCAGTGCACAGAGCAGCCACGCTGGACAGCAACTCAGTCAGAAATGGGGACAAG 1740
Qy 1876 CTCCAAAGAGACCTGAGAAGCGCAGCAGTCCCTTGGGGCTTCTTTAAAGGCTCGGA 1935
Db 1741 CTCCAAAGAGACCTGAGAAGCGCAGCAGTCCCTTGGGGCTTCTTTAAAGGCTCGGA 1800
Qy 1936 CCAAAGCGATGTTGGATGCTCAAGTGCAACAGACCCAGTATCATCGGACCAAGTTGGC 1995
Db 1801 CCAAAGCGATGTTGGATGCTCAAGTGCAACAGACCCAGTATCATCGGACCAAGTTGGC 1860
Qy 1996 AARCCCAAGTAAACAATCAGCAGCGTTCCACAGGTTCTCCGCCACCAAGATGTT 2055
Db 1861 AARCCCAAGTAAACAATCAGCAGCGTTCCACAGGTTCTCCGCCACCAAGATGTT 1920
Qy 2056 CTCCTTACTCCATCTCCTCCCAACACAGCTCCATGTATATATTTCTGTATGCCAGCA 2115
Db 1921 CTCCTTACTCCATCTCCTCCCAACACAGCTCCATGTATATATTTCTGTATGCCAGCA 1980
Qy 2116 AATGAATTTGCTAGAAATTAAGCCGAGCTGTTGTATATTTAGGTGTAATTTACG 2175
Db 1981 AATGAATTTGCTAGAAATTAAGCCGAGCTGTTGTATATTTAGGTGTAATTTACG 2040
Qy 2176 TCTCTGCTCAGTCTTTCTGCAATAAACAAGTAAAGATGTTTACAGGTCACCTAGTT 2235
Db 2041 TCTCTGCTCAGTCTTTCTGCAATAAACAAGTAAAGATGTTTACAGGTCACCTAGTT 2100
Qy 2236 GGGTCAGAAGAGTCGATGATCACAAGCAGGAGGAGGGAATAGAGGAATGTTTCGG 2295
Db 2101 GGGTCAGAAGAGTCGATGATCACAAGCAGGAGGAGGGAATAGAGGAATGTTTCGG 2160
Qy 2296 GTTAAGTATGAATAATGTCAGTGTGGCCGGCGTGTGGCTCTCGCTGTAAATCTCAGC 2355
Db 2161 GTTAAGTATGAATAATGTCAGTGTGGCCGGCGTGTGGCTCTCGCTGTAAATCTCAGC 2220
Qy 2356 ACTTTGGAGCCGAGGCGAGTGATCACCCTGAGTTCAGAGTTCAGACTAGCTCGCC 2415
Db 2221 ACTTTGGAGCCGAGGCGAGTGATCACCCTGAGTTCAGAGTTCAGACTAGCTCGCC 2280
Qy 2416 AACATCATGAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGGCGATGGTGGCACAC 2475
Db 2281 AACATCATGAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGGCGATGGTGGCACAC 2340
Qy 2476 ACCTGTAGTCCAGTACTCGGGAGCCCAACGACAGAACCGTTGTACCCAGGAGTG 2535
Db 2341 ACCTGTAGTCCAGTACTCGGGAGCCCAACGACAGAACCGTTGTACCCAGGAGTG 2400
Qy 2536 GAGGTTGAGTGAGCCGAGTTGACCATTTGCATCTCCACCTGGGCGACAGCAAGATT 2595
Db 2401 GAGGTTGAGTGAGCCGAGTTGACCATTTGCATCTCCACCTGGGCGACAGCAAGATT 2460
Qy 2596 CTATCAAAAAAAGGAGGAGTGAAGTATAGAGGAATGCTGTAGAAGGAA 2655
Db 2461 CTATCAAAAAAAGGAGGAGTGAAGTATAGAGGAATGCTGTAGAAGGAA 2520
Qy 2656 TTAAGCGTTGATGAACCGTGTCTATCTCTAAGCTTGAAGAGGAGAGCAAAATCC 2715
Db 2521 TTAAGCGTTGATGAACCGTGTCTATCTCTAAGCTTGAAGAGGAGAGCAAAATCC 2580
Qy 2716 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTGGTGTGGTCCCAAT 2775
Db 2581 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTGGTGTGGTCCCAAT 2640
Qy 2776 TTCTTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTTCAAT 2835
Db 2641 TTCTTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTTCAAT 2700
Qy 2836 AGGTCCTTGAAGTTATGATGAGAGGAGAAAGATTGACTCGGGAGGGCTTAAATGAT 2895
Db 2701 AGGTCCTTGAAGTTATGATGAGAGGAGAAAGATTGACTCGGGAGGGCTTAAATGAT 2760

Qy 2896 TTGGGAAAAAATTTGCTTTTGGGCTCAGTGACAAACGGCAAAAGATTACAACTT 2948
Db 2761 TTGGGAAAAAATTTGCTTTTGGGCTCAGTGACAAACGGCAAAAGATTACAACTT 2813

RESULT 6
ADQ18993

ID ADQ18993 standard; DNA; 2813 BP.

XX ADQ18993;

XX 26-AUG-2004 (first entry)

XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 1812.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.

OS Homo sapiens.

XX W02004048938-A2.

XX 10-JUN-2004.

XX 26-NOV-2003; 2003WO-US038193.

XX 26-NOV-2002; 2002US-0429739P.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Aziz N, Ginsburg WM, Zlotnik A;

XX WPI; 2004-441208/41.

XX Early detection of soft tissue sarcoma comprises determining expression
of a gene in a first soft tissue sample and a normal soft tissue sample
and comparing the gene expression, also useful in treating soft tissue
sarcoma.

XX Example 2; SEQ ID NO 1812; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma
which comprises obtaining a first soft tissue sample from an individual
and a normal soft tissue sample from the same or different individual,
determining the expression of a gene in both samples and comparing the
expression of the gene in both soft tissue samples, where a higher level
of protein expression in the first soft tissue sample indicates the
presence of soft tissue sarcoma. The method of the invention has
cytostatic applications and may be useful for detecting soft tissue
sarcoma, possibly via gene therapy or vaccine production. The nucleic
acid sequences may be useful in diagnostic and screening applications.
The current sequence is that of a human soft tissue sarcoma-upregulated
DNA of the invention. The current sequence is not shown within the
specification per se but was submitted in CD format by the inventor.

XX Sequence 2813 BP; 868 A; 671 C; 758 G; 516 T; 0 U; 0 Other;

Query Match 91.7%; Score 2813; DB 12; Length 2813;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 CTCACCCGACAGCCAGGCACTGGGCGACGACGACTGGAGACCCAGGACCCCTGTGCAGGAG 195
Db 1 CTCACCCGACAGCCAGGCACTGGGCGACGACGACTGGAGACCCAGGACCCCTGTGCAGGAG 60

Qy 196 CAGTCCCGGTGACACAGGGGAGCTGAAGATCTCCACAGGGGCTCAGCAGGAGCAATG 255
Db 61 CAGTCCCGGTGACACAGGGGAGCTGAAGATCTCCACAGGGGCTCAGCAGGAGCAATG 120

Qy 256 GGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAGACCAAGAGATGAACCAAGAGCAGAG 315
Db 121 GGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAGACCAAGAGATGAACCAAGAGCAGAG 180

QY 316 ACTTACCAGGCAACCGCTCTGCTCTGAACGGGGTTCCAGTGGTGGTGTGACCCACACA 375
DB |||||
DB 181 ACTTACCAGGCAACCGCTCTGCTCTGAACGGGGTTCCAGTGGTGGTGTGACCCACACA 240
QY 376 GTTCAGCACTTAGAGGAAGTGCATTTGGGAATAAGTGTCAAGACGGATAAATGTGGCCACT 435
DB |||||
DB 241 GTTCAGCACTTAGAGGAAGTGCATTTGGGAATAAGTGTCAAGACGGATAAATGTGGCCACT 300
QY 436 TCTTCCCGGAGACAACCGAGAGATAGTGTGTGCGGATGCAACCGGAAGAATCTTGGG 495
DB |||||
DB 301 TCTTCCCGGAGACAACCGAGATAAGTGTGTGCGGATGCAACCGGAAGAATCTTGGG 360
QY 496 AAAGGSCCAACCCGAGGCCACAGCTGCTTAATCTCGTTTTTTTCTTGATGCTCTCTCGG 555
DB |||||
DB 361 AAAGGSCCAACCCGAGGCCACAGCTGCTTAATCTCGTTTTTTTCTTGATGCTCTCTCGG 420
QY 556 CCTGTACAGGACGTACCGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTT 615
DB |||||
DB 421 CCTGTACAGGACGTACCGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTT 480
QY 616 GATGTCAAGTCCAAATAAGTCCAGCGAAACAAAGACCCCAAGTGAGCTGACACTTCCG 675
DB |||||
DB 481 GATGTCAAGTCCAAATAAGTCCAGCGAAACAAAGACCCCAAGTGAGCTGACACTTCCG 540
QY 676 GTGGCAGCTGACCGGGCAGGACACAGATAAACCCGAGGCCACGCCCGCCGCCAAGAC 735
DB |||||
DB 541 GTGGCAGCTGACCGGGCAGGACACAGATAAACCCGAGGCCACGCCCGCCGCCAAGAC 600
QY 736 AAGGTCTCTCTGCGCCGAGGATCCACAGCTTCTCCACCTTGAGACAGGGGGAGCAGGA 795
DB |||||
DB 601 AAGGTCTCTCTGCGCCGAGGATCCACAGCTTCTCCACCTTGAGACAGGGGGAGCAGGA 660
QY 796 GGAGAGCTCCCTCCAAGCCCAAGGACTCCAGCTTTTTTTTGACAAATCTTCAAGCTGGAC 855
DB |||||
DB 661 GGAGAGCTCCCTCCAAGCCCAAGGACTCCAGCTTTTTTTTGACAAATCTTCAAGCTGGAC 720
QY 856 AAGGACAGGAATAAGGTCCAGGTGACAGCCAAAGGAAGCCAGAGGGCAGAGATCAA 915
DB |||||
DB 721 AAGGACAGGAATAAGGTCCAGGTGACAGCCAAAGGAAGCCAGAGGGCAGAGATCAA 780
QY 916 GACAGGTGGATGAGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCTGCGAGGAG 975
DB |||||
DB 781 GACAGGTGGATGAGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCTGCGAGGAG 840
QY 976 GACATAGTTGACGCAAGGAAAAAGAGGACAAAGAACTTGGAACTGCGGATTTGCTCTGTC 1035
DB |||||
DB 841 GACATAGTTGACGCAAGGAAAAAGAGGACAAAGAACTTGGAACTGCGGATTTGCTCTGTC 900
QY 1036 CTTGGGACCCAGAAAGGACTGAGACTGCAAGGACGATTCCTGAGGACAGCAGCTATAGCA 1095
DB |||||
DB 901 CTTGGGACCCAGAAAGGACTGAGACTGCAAGGACGATTCCTGAGGACAGCAGCTATAGCA 960
QY 1096 GAGATATAATTCATCATGATTTCTTTAAACTCTGTTTCACTTAAACAAAGCTGAA 1155
DB |||||
DB 961 GAGATATAATTCATCATGATTTCTTTAAACTCTGTTTCACTTAAACAAAGCTGAA 1020
QY 1156 ACAAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACACTTCAGCTGACCTT 1215
DB |||||
DB 1021 ACAAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACACTTCAGCTGACCTT 1080
QY 1216 AAGTCAGACAAAGCCAACTTTATCCAGAGACCCAAAGGGGTGGCAAGAAATTCACAA 1275
DB |||||
DB 1081 AAGTCAGACAAAGCCAACTTTATACATCCAGAGACCCAAAGGGGTGGCAAGAAATTCACAA 1140
QY 1276 GGATGCAACCATCGGGGACACACAGTCCGTGACAACTTGAACCTGACGAGGAGGC 1335
DB |||||
DB 1141 GGATGCAACCATCGGGGACACACAGTCCGTGACAACTTGAACCTGACGAGGAGGC 1200
QY 1336 ACCAAGGAGAAATCAGGACCCACTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1395
DB |||||
DB 1201 ACCAAGGAGAAATCAGGACCCACTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1260

QY 1396 GTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGAGATGTGTTGTGAGTCCACAGTA 1455
DB |||||
DB 1261 GTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGAGATGTGTTGTGAGTCCACAGTA 1320
QY 1456 GAGATTATAAGTCCAAAGGAAGTGAATCAGCTTTACAAACAGTGCAGCTCAACGAAGGA 1515
DB |||||
DB 1321 GAGATTATAAGTCCAAAGGAAGTGAATCAGCTTTACAAACAGTGCAGCTCAACGAAGGA 1380
QY 1516 GATGTGCACCTGAACCCACAGAAAGCGAAACTCAAAAGAGAGAAAGCAAAACCAAGAAAC 1575
DB |||||
DB 1381 GATGTGCACCTGAACCCACAGAAAGCGAAACTCAAAAGAGAGAAAGCAAAACCAAGAAAC 1440
QY 1576 TCTCTGATGSGGTTTCTCAGACCAATGTCTAGTGAAGGGGATGAGGGATCACCCACTCA 1635
DB |||||
DB 1441 TCTCTGATGSGGTTTCTCAGACCAATGTCTAGTGAAGGGGATGAGGGATCACCCACTCA 1500
QY 1636 GAAGAAATAATGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGAGACTATC 1695
DB |||||
DB 1501 GAAGAAATAATGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGAGACTATC 1560
QY 1696 ACACCGCAGAGCTGAACCAACAGGAGCACCAAGAGGGTAAAGGGGCTCTCTCGAAG 1755
DB |||||
DB 1561 ACACCGCAGAGCTGAACCAACAGGAGCACCAAGAGGGTAAAGAGGGCTCTCTCGAAG 1620
QY 1756 GACAAAGTCAAGCAGCGGAGATCAACAGCAGAGAGCAACAGCAGGAGAGCCAAAGAA 1815
DB |||||
DB 1621 GACAAAGTCAAGCAGCGGAGATCAACAGCAGAGAGCAACAGCAGGAGAGCCAAAGAA 1680
QY 1816 CCAGCCAGTGCACAGAGCAGGCCACAGTGCAGACGAACTCACCTGCAGATGGGAGCAAG 1875
DB |||||
DB 1581 CCAGCCAGTGCACAGAGCAGGCCACAGTGCAGACGAACTCACCTGCAGATGGGAGCAAG 1740
QY 1876 CTCCAAAGAGACTGAGAAAGCGGACAGTCCCTTGGGGGCTCTTTAAAGGCTTGGGA 1935
DB |||||
DB 1741 CTCCAAAGAGACTGAGAAAGCGGACAGTCCCTTGGGGGCTCTTTAAAGGCTTGGGA 1800
QY 1936 CCAAGCGGATGTTGGATGCTCAAGTGCACACAGCCAGTATCATCGGACAGTGGC 1995
DB |||||
DB 1801 CCAAGCGGATGTTGGATGCTCAAGTGCACACAGCCAGTATCCATCGGACAGTGGC 1860
QY 1996 AAACCCAAAGTAAACAAATCAGACGGTTCCACACAGGTTCTCCTGCCACCAAGATGTGT 2055
DB |||||
DB 1861 AAACCCAAAGTAAACAAATCAGACGGTTCCACACAGGTTCTCCTGCCACCAAGATGTGT 1920
QY 2056 CTCTTACTCCATCTCCTCCCAACACAGCTCCATATATATATTTCTCTGATGGCCAGCA 2115
DB |||||
DB 1921 CTCTTACTCCATCTCCTCCCAACACAGCTCCATATATATATTTCTCTGATGGCCAGCA 1980
QY 2116 AATGAAATTCGCTAGAAATTAAGCCGAGCTGTTGTATATTTAGAGTGTATTTATAG 2175
DB |||||
DB 1981 AATGAAATTCGCTAGAAATTAAGCCGAGCTGTTGTATATTTAGAGTGTATTTATAG 2040
QY 2176 TCTCTGTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTAGCAGGTCACTAGTT 2235
DB |||||
DB 2041 TCTCTGTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTAGCAGGTCACTAGTT 2100
QY 2236 GGGTCAAGAGTGCATGATCAACCAAGCAGGAAAGGGAGGAAATAGAGGAATGTGTTCCG 2295
DB |||||
DB 2101 GGGTCAAGAGTGCATGATCAACCAAGCAGGAAAGGGAGGAAATAGAGGAATGTGTTCCG 2160
QY 2296 GTTAAAGTGAATAATGGCAGTGGTGGCCGGCTGTGGCTCTCGCTGTGTAATCTCAGC 2355
DB |||||
DB 2161 GTTAAAGTGAATAATGGCAGTGGTGGCCGGCTGTGGCTCTCGCTGTGTAATCTCAGC 2220
QY 2356 ACTTTGGAGGCGCAGGACAGTGGATCACTGAGTTCAGAGTTTCAGAGCTAGCTGGCC 2415
DB |||||
DB 2221 ACTTTGGAGGCGCAGGACAGTGGATCACTGAGTTCAGAGTTTCAGAGCTAGCTGGCC 2280
QY 2416 AACATCATGAAACCCCTCTCTACTTAAATAATACAAAAATTAGCCAGGCTGGTGGCACAC 2475
DB |||||
DB 2281 AACATCATGAAACCCCTCTCTACTTAAATAATACAAAAATTAGCCAGGCTGGTGGCACAC 2340
QY 2476 ACCTGTAGTCCAGCTACTCTGGGAGCCCAACGCAACGCAACCCGCTTGTATCCAGGAGGTG 2535

[illegible]

RESULT 7

ADR25076
ID ADR25076 standard; DNA; 2813 BP.

AC ADR25076:

21-OCT-2004 (first entry)

DE Breast cancer prognosis marker #937.

ds: breast cancer; prognosis; gene expression; diagnosis.

OS Homo sapiens.

AA
PN
WO2004065545-A2.XX
PD 05-AUG-2004.

15-JAN-2004; 2004WO-US001100.

XX 15-JAN-2003: 2003US-00342887. PR

XX PA (ROSE-) ROSETTA INPHARMATICS LLC.

PA (NECA-) NETHERLANDS CANCER INST.

PI Van't Veer LJ, He Y;

DR WPI: 2004-593473/57.

Classifying a breast cancer patient according to prognosis comprises determining the similarity between the level of expression of each of five genes in a cell sample taken from patient, to control levels.

PS Disclosure; SEQ ID NO 937; 226pp; English.

The invention relates to a method of classifying a breast cancer patient according to prognosis by determining the similarity between the level of expression of each of five genes for which markers are listed in the specification, in a cell sample taken from the breast cancer patient, to control levels of expression for each respective five genes to obtain a patient similarity value. The methods are useful for classifying a breast

CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.

Sequence 2813 BP; 868 A; 671 C; 758 G; 516 T; 0 U; 0 Other;

| | | | | |
|-----------------------|--------------|--------------|------------|--------------|
| Query Match | 91.7% | Score 2813; | DB 13; | Length 2813; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 2813; | Conservative | 0; | Mismatches | 0; |
| Indels | 0; | | | Gaps |

| | | | |
|----|------|--|------|
| QY | 136 | CTCACCCAGCAGCCAGGCACTGGGCAGCACGCACTGGAGAGCCAGAGACCCTGTGCAGGAG | 195 |
| DB | 1 | CTCCACCGCAGCCAGGCACTGGGCAGCACGCACTGGAGAGCCAGGACCCCTGTGCAGGAG | 60 |
| QY | 196 | CAGCTCCGGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCAGAGCAATG | 255 |
| DB | 61 | CAGCTCCGGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCAGAGCAATG | 120 |
| QY | 256 | GGTAAACCAATAGTGTTCGCCAAGAAGTTGAAGACCAAGAGAAATGAACCCAGAAAGCAGAG | 315 |
| DB | 121 | GGTAAACCAATAGTGTTCGCCAAGAAGTTGAAGACCAAGAGAAATGAACCCAGAAAGCAGAG | 180 |
| QY | 316 | ACTTACCAAGCAACCGCTCTGCTCTGAAACGGGGTTCACGTGTGGTGTGCAGCCACACA | 375 |
| DB | 181 | ACTTACCAAGCAACCGCTCTGCTCTGAAACGGGGTTCACGTGTGGTGTGCAGCCACACA | 240 |
| QY | 376 | GTTCAGCACTTAGAGGAAGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT | 435 |
| DB | 241 | GTTCAGCACTTAGAGGAAGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT | 300 |
| QY | 436 | TCCTTCCCGGAGACAAACGGAGATAAGTGTGTGGGATGCCAAACGGAAAGAAATCTTTGGG | 495 |
| DB | 301 | TCCTTCCCGGAGACAAACGGAGATAAGTGTGTGGGATGCCAAACGGAAAGAAATCTTTGGG | 360 |
| QY | 496 | AAAGAGGCCAAACCCGAGGACACAGCTGCTAAATCTCGTTTTTCTTGATGCTCTCTCGG | 555 |
| DB | 361 | AAAGAGGCCAAACCCGAGGACACAGCTGCTAAATCTCGTTTTTCTTGATGCTCTCTCGG | 420 |
| QY | 556 | CCTGTACAGGACGTACCGGAGACCAAGCCGACAGATTATCCCTTGGATCAGTGAAGCTT | 615 |
| DB | 421 | CCTGTACAGGACGTACCGGAGACCAAGCCGACAGATTATCCCTTGGATCAGTGAAGCTT | 480 |
| QY | 616 | GATGTACGTCCCAATAAGCTCAGCGAAACAAAGACCCAAAGTCAGAGCTGGAACATTCG | 675 |
| DB | 481 | GATGTACGTCCCAATAAGCTCAGCGAAACAAAGACCCAAAGTCAGAGCTGGAACATTCG | 540 |
| QY | 676 | GTGCGAGCTGACCGGGGAGGACACAGATAAAACCCAGGGACGCCCCGGGCCCAAGAC | 735 |
| DB | 541 | GTGCGAGCTGACCGGGGAGGACACAGATAAAACCCAGGGACGCCCCGGGCCCAAGAC | 600 |
| QY | 736 | AAGGTCTCTCTGCGCCAGGGATCCACGCTTCTCCACCTGAGACAGGGGGAGCAGGA | 795 |
| DB | 601 | AAGGTCTCTCTGCGCCAGGGATCCACGCTTCTCCACCTGAGACAGGGGGAGCAGGA | 660 |
| QY | 796 | GGAGAGCTCCCTCCAAGCCCAAGGACTCCAGCTTTTTTTCAGAAATTTCTTCAAGCTGGAC | 855 |
| DB | 661 | GGAGAGCTCCCTCCAAGCCCAAGGACTCCAGCTTTTTTTCAGAAATTTCTTCAAGCTGGAC | 720 |
| QY | 856 | AAGGGACAGGAAAAGGTGCCAGGTGACAGCCAAACAGGAAGCCAAAGGGGACAGACATCAA | 915 |
| DB | 721 | AAGGGACAGGAAAAGGTGCCAGGTGACAGCCAAACAGGAAGCCAAAGGGGACAGACATCAA | 780 |
| QY | 916 | GACAAAGTGGATGAGTTCTCTGGCTTTATCAGGGCAGTCCGATGATGTCCTGTGACAGGAAG | 975 |
| DB | 781 | GACAAAGTGGATGAGTTCTCTGGCTTTATCAGGGCAGTCCGATGATGTCCTGTGACAGGAAG | 840 |
| QY | 976 | GACATAGTTGACGCGAAGGAAAAGAGGACAGAACTTCGAACTGCGGATGCTCTGTC | 1035 |
| DB | 841 | GACATAGTTGACGCGAAGGAAAAGAGGACAGAACTTCGAACTGCGGATGCTCTGTC | 900 |
| QY | 1036 | CCTGGGACCCAGAAAGACTGGAGACTGCAAGAGGACGATTTCCAGGGCAGCAGCTATAGCA | 1095 |

Db 901 CCTGGGACCCAGAGGACTGGAGACTGCAAGAGACGATTCACAGGACGAGCTATAGCA 960
Qy 1096 GAGATATATATTCATCATGAGTTCTTTAAACCTCTGGTTTACCTAAACAAGCTGAA 1155
Db 961 GAGATATATATTCATCATGAGTTCTTTAAACCTCTGGTTTACCTAAACAAGCTGAA 1020
Qy 1156 ACAAAAAAGGACCCAGAGAGACGCGGTCTGAAAAAGTCAACCACACCTTCAGCTGACCTT 1215
Db 1021 ACANAAAGGACCCAGAGAGACGCGGTCTGAAAAAGTCAACCACACCTTCAGCTGACCTT 1080
Qy 1216 AAGTCAGACAAAGCCAACTTTACATPCCAGAGAGACCCAAAGGCGCTGGCAAGAATTCCTAAA 1275
Db 1081 AAGTCAGACAAAGCCAACTTTACATPCCAGAGAGACCCAAAGGCGCTGGCAAGAATTCCTAAA 1140
Qy 1276 GGATGCAACCCATCGGGGACACACAGTCCGTGCAACCCCTGAACTTCGCAAGGAGGC 1335
Db 1141 GGATGCAACCCATCGGGGACACACAGTCCGTGCAACCCCTGAACTTCGCAAGGAGGC 1200
Qy 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTGTTGGAAAAAGTCA 1395
Db 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTGTTGGAAAAAGTCA 1260
Qy 1396 GTTAAAGAGGACTCAGTCCCAAGGTCGCGAGGAGAAATGTGGTGTGTGAGTCAACAGTA 1455
Db 1261 GTTAAAGAGGACTCAGTCCCAAGGTCGCGAGGAGAAATGTGGTGTGTGAGTCAACAGTA 1320
Qy 1456 GAGATTATTAAGTCAAGGAAGTAGAATCAGCTTTAACAACAGTGGAGCTCAACGAAGGA 1515
Db 1321 GAGATTATTAAGTCAAGGAAGTAGAATCAGCTTTAACAACAGTGGAGCTCAACGAAGGA 1380
Qy 1516 GATGCTGACCTGAAACCCACAGAGCGGAACTCTCAAAAGAGAGAAAGCAACACCAAGAAC 1575
Db 1381 GATGCTGACCTGAAACCCACAGAGCGGAACTCTCAAAAGAGAGAAAGCAACACCAAGAAC 1440
Qy 1576 TCTCTGATGGGTTTCTCAGACAAATGTCAAGTAAAGGGGATGGAGGATCACCCACTCA 1635
Db 1441 TCTCTGATGGGTTTCTCAGACAAATGTCAAGTAAAGGGGATGGAGGATCACCCACTCA 1500
Qy 1636 GAAGAAATTAATGGGAAGACTCCAGCTGCGCAACATCAGACTCAACAGAAAGACTATC 1695
Db 1501 GAAGAAATTAATGGGAAGACTCCAGCTGCGCAACATCAGACTCAACAGAAAGACTATC 1560
Qy 1696 ACACCGCAGAGCTGAACCAACAGGAGCACCACAGAGGGTAAAGGGGCTCTCCGAAG 1755
Db 1561 ACACCGCAGAGCTGAACCAACAGGAGCACCACAGAGGGTAAAGGGGCTCTCCGAAG 1620
Qy 1756 GACAAGAGTCAAGAGCCGAGATGAACAGAGAGAGCAACAGAGGAGAGCAAGCAAGAA 1815
Db 1621 GACAAGAGTCAAGAGCCGAGATGAACAGAGAGAGCAACAGAGGAGAGCAAGCAAGAA 1680
Qy 1816 CCAGCCAGTGCACAGAGCAGGCCACGCTGAGACAGCAACTCACTGCAAGATGGGGACAAG 1875
Db 1681 CCAGCCAGTGCACAGAGCAGGCCACGCTGAGACAGCAACTCACTGCAAGATGGGGACAAG 1740
Qy 1876 CTCCTAAAGAGACTGAGAGCGGACGAGTCCCTTGGGGCTCTTTAAAGGCTCGGA 1935
Db 1741 CTCCTAAAGAGACTGAGAGCGGACGAGTCCCTTGGGGCTCTTTAAAGGCTCGGA 1800
Qy 1936 CCAAGCGGATGTTGGATGCTCAAGTGCACAGAGCAGAGTATCCATCGGACCAAGTTGGC 1995
Db 1801 CCAAGCGGATGTTGGATGCTCAAGTGCACAGAGCAGAGTATCCATCGGACCAAGTTGGC 1860
Qy 1996 AAAACCAAGTAAACAAATCAGACGCGTTCCACAGGTTCTCTGCCACCAAGATGTGT 2055
Db 1861 AAAACCAAGTAAACAAATCAGACGCGTTCCACAGGTTCTCTGCCACCAAGATGTGT 1920
Qy 2056 CTCCTTACTCTCTCTCCCAACAGCTCCATGATATATTTCTTCTGATGCCAGCA 2115
Db 1921 CTCCTTACTCTCTCTCCCAACAGCTCCATGATATATTTCTTCTGATGCCAGCA 1980
Qy 2116 AATGAAATTCGCTTAGAAAATTAAGCCCGAGCTGTGTTATATTCAGGTGTTATTTACG 2175
Db 1981 AATGAAATTCGCTTAGAAAATTAAGCCCGAGCTGTGTTATATTCAGGTGTTATTTACG 2040

Qy 2176 TCTCTGGTCCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTTAGCAGGTCACCTAGTT 2235
Db 2041 TCTCTGGTCCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTTAGCAGGTCACCTAGTT 2100
Qy 2236 GGGTCAGAGAGTGCATGATCACCACAGCAGGAAAGGAGGAATAGAGGAATGTGTTCCG 2295
Db 2101 GGGTCAGAGAGTGCATGATCACCACAGCAGGAAAGGAGGAATAGAGGAATGTGTTCCG 2160
Qy 2296 GTTAAAGTGAATAAATGGCAGTGGTGGCCGGCGCTGTGGCTCTCGCTCTGAATCTCAGC 2355
Db 2161 GTTAAAGTGAATAAATGGCAGTGGTGGCCGGCGCTGTGGCTCTCGCTCTGAATCTCAGC 2220
Qy 2356 ACTTTGGAGGCCAGCAGCAGTGGATCACCTGAGGTCAGGAGTTCAAGACTAGCCTGGCC 2415
Db 2221 ACTTTGGAGGCCAGCAGCAGTGGATCACCTGAGGTCAGGAGTTCAAGACTAGCCTGGCC 2280
Qy 2416 AACATCATGAAACCCCTCTCTACTAAATAATACAAAAATTAGCCAGGCTGTCGCACAC 2475
Db 2281 AACATCATGAAACCCCTCTCTACTAAATAATACAAAAATTAGCCAGGCTGTCGCACAC 2340
Qy 2476 ACTGTAGTCCAGCTACTCGGGAGCCCAACGCAACGAGAACCGCTTGTACCCAGGAGTG 2535
Db 2341 ACTGTAGTCCAGCTACTCGGGAGCCCAACGCAACGAGAACCGCTTGTACCCAGGAGTG 2400
Qy 2536 GAGTTGAGTGGCCGAAAGTTGCACTTGCATCTCCACCTCGGCGCAGCAGCAAGATT 2595
Db 2401 GAGTTGAGTGGCCGAAAGTTGCACTTGCATCTCCACCTCGGCGCAGCAGCAAGATT 2460
Qy 2596 CTATCAAAAAAAGGAGCAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGCTAGAGGAA 2655
Db 2461 CTATCAAAAAAAGGAGCAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGCTAGAGGAA 2520
Qy 2656 TTAAGCGTTGTAGTAAACGCGTCTCATCTCTTAAGCTTGAAGAGGAGAGCAAAAAATCC 2715
Db 2521 TTAAGCGTTGTAGTAAACGCGTCTCATCTCTTAAGCTTGAAGAGGAGAGCAAAAAATCC 2580
Qy 2716 ATTTGTTTAAATTCACATCTCAGGAGGAGAGAACCCGGGCTGTGTTGGTGGTTCCTCAT 2775
Db 2581 ATTTGTTTAAATTCACATCTCAGGAGGAGAGAACCCGGGCTGTGTTGGTGGTTCCTCAT 2640
Qy 2776 TTCTAGAACCGAATGTGGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTTCAAT 2835
Db 2641 TTCTAGAACCGAATGTGGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTTCAAT 2700
Qy 2836 AGGTCTCTGTAAAGTTATGATGAGAGGAGAGAAAGATTGACTGGGGAGGCTTAAATCAT 2895
Db 2701 AGGTCTCTGTAAAGTTATGATGAGAGGAGAGAAAGATTGACTGGGGAGGCTTAAATCAT 2760
Qy 2896 TTGGGAAAAAATTTGCTTTTGGGCTCAGTGCACACCGGCAAGATTACACTT 2948
Db 2761 TTGGGAAAAAATTTGCTTTTGGGCTCAGTGCACACCGGCAAGATTACACTT 2813

RESULT 8

ACC90605

ID ACC90605 standard; DNA; 2020 BP.

XX ACC90605;

AC ACC90605;

XX 12-AUG-2003 (first entry)

DT 12-AUG-2003 (first entry)

XX Human CGDP-28 encoding DNA.

DE Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

| | | | | | |
|----|--|----|------|--|------|
| XX | 03-APR-2003. | Db | 126 | GGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGATGAACCA | 185 |
| XX | 26-SEP-2002; 2002WO-US031095. | Qy | 307 | GAAGCAGACAGACTTACCAGGACAACCGCTCTGCTCTGAACGGGGTTCCAGTGGTGTGCG | 366 |
| XX | 28-SEP-2001; 2001US-0326389P. | Db | 186 | GAAGCAGACAGACTTACCAGGACAACCGCTCTGCTCTGAACGGGGTTCCAGTGGTGTGCG | 245 |
| XX | 05-OCT-2001; 2001US-0327380P. | Qy | 367 | ACCCACACAGTTTACGACCTTAGAGGAAGTGCAGCTTGGGAATAAGTGTCAAGACGGATAAT | 426 |
| XX | 12-OCT-2001; 2001US-0329690P. | Db | 246 | ACCCACACAGTTTACGACCTTAGAGGAAGTGCAGCTTGGGAATAAGTGTCAAGACGGATAAT | 305 |
| XX | 26-OCT-2001; 2001US-0345384P. | Qy | 427 | GTGGCCACTTCTTCCCGGAGACAACGGAGATAAGTGTCTGCGGATGCGCAACGGAAAG | 486 |
| XX | 02-NOV-2001; 2001US-0350219P. | Db | 306 | GTGGCCACTTCTTCCCGGAGACAACGGAGATAAGTGTCTGCGGATGCGCAACGGAAAG | 365 |
| XX | 09-NOV-2001; 2001US-0344518P. | Qy | 487 | AATCTTGGGAAAGAGGCCAAACCCGAGGCCACAGCTGCTAAATCTCGTTTTTCTTGATG | 546 |
| XX | 16-NOV-2001; 2001US-0345143P. | Db | 366 | AATCTTGGGAAAGAGGCCAAACCCGAGGCCACAGCTGCTAAATCTCGTTTTTCTTGATG | 425 |
| XX | 03-DEC-2001; 2001US-0336908P. | Qy | 547 | CTCTCTCGGCTGTACCCAGGACGTACCGGAGACAACAGCCGACAGATTATCCCTTGGATCA | 606 |
| XX | 07-DEC-2001; 2001US-0340747P. | Db | 426 | CTCTCTCGGCTGTACCCAGGACGTACCGGAGACAACAGCCGACAGATTATCCCTTGGATCA | 485 |
| XX | (INCY-) INCYTE GENOMICS INC. | Qy | 607 | GTGAAGCTTATGTCTCAGCTCCATTAAGCTCCAGCGCAACAAAGACCAAGCTGAGAGCTGG | 666 |
| XX | Azimzai Y, Baughn MR, Becha SD, Borowsky ML, Chawla NK; | Db | 545 | GTGAAGCTTATGTCTCAGCTCCATTAAGCTCCAGCGCAACAAAGACCAAGCTGAGAGCTGG | 545 |
| XX | Elliott VS, Emerling BM, Gandhi AR, Gietzen KJ, Gorvad AE; | Qy | 667 | ACACTTCCGCTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCAGGACGACGCCCGG | 726 |
| XX | Griffin JA, Hafalia AJA, Ison CH, Kable AE, Kalafus DP; | Db | 546 | ACACTTCCGCTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCAGGACGACGCCCGG | 605 |
| XX | Lehr-Mason PM, Lu DAM, Marquis JP, Nguyen DB, Rankumar J; | Qy | 727 | GCCCAAGACAAGTCTCTCTGCGGCCAGGGATCCCAAGCTTCTCCACCTGAGACAGGG | 786 |
| XX | Richardson TW, Sapperstein SK, Swarnakar A, Tang YT, Tran UK; | Db | 606 | GCCCAAGACAAGTCTCTCTGCGGCCAGGGATCCCAAGCTTCTCCACCTGAGACAGGG | 665 |
| XX | Warren BA, Xu Y, Yao MG, Yue H, Yue H; | Qy | 787 | GGAGCAGGAGGAGAGTCCCTCCAGCCCAAGGACTCCAGCTTTTGTGCAAAATTTCTTC | 846 |
| XX | WPI; 2003-421159/39. | Db | 666 | GGAGCAGGAGGAGAGTCCCTCCAGCCCAAGGACTCCAGCTTTTGTGCAAAATTTCTTC | 725 |
| XX | P-PSDB; ABR69628. | Qy | 847 | AAGCTGCAACAGGACAGGAAAAGTCCAGGTGACAGCCAAACAGGAAGCAAGAGGGCA | 906 |
| XX | New human proteins associated with cell growth, differentiation, and | Db | 726 | AAGCTGCAACAGGACAGGAAAAGTCCAGGTGACAGCCAAACAGGAAGCAAGAGGGCA | 785 |
| XX | death (CGDD), useful for diagnosing, treating and preventing diseases or | Qy | 907 | GAGCATCAAGACAAGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCT | 966 |
| XX | conditions associated with the aberrant CGDD expression e.g. cancer, | Db | 786 | GAGCATCAAGACAAGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCT | 845 |
| XX | AIDS, or epilepsy. | Qy | 967 | GCAGGGAAGACATAGTTGACGCAAGCAAGAAAAGAGCAAGAACTTGGAACTGCCGAT | 1026 |
| XX | Claim 5; Page 330; 350pp; English. | Db | 846 | GCAGGGAAGACATAGTTGACGCAAGCAAGAAAAGAGCAAGAACTTGGAACTGCCGAT | 905 |
| XX | The invention relates to an isolated polypeptide associated with cell | Qy | 1027 | TGCTCTCTCTGGGACCCAGAGGACTGGAGACTGCAAGGACGATTCCAGGCGACGA | 1086 |
| XX | growth, differentiation and death (CGDD). Also disclosed are the | Db | 906 | TGCTCTCTCTGGGACCCAGAGGACTGGAGACTGCAAGGACGATTCCAGGCGACGA | 965 |
| XX | polynucleotides encoding the polypeptides. The polypeptides and | Qy | 1087 | GCTATAGCAGAGATAATAATTCATCATGATTTTCTTTAAACTCTGGTTTCACTTAAC | 1146 |
| XX | polynucleotides are useful in diagnosing, treating and preventing or | Db | 966 | GCTATAGCAGAGATAATAATTCATCATGATTTTCTTTAAACTCTGGTTTCACTTAAC | 1025 |
| XX | diseases or conditions associated with the decreased expression or | Qy | 1147 | AAAGCTGAAACAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACCTTCA | 1206 |
| XX | expression of CGDD. Such diseases include cell proliferative (e.g. | Db | 1026 | AAAGCTGAAACAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACCTTCA | 1085 |
| XX | cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's | Qy | 1207 | GCTGACCTTAAGTCAGACAAGCCAACTTTATATCCAGGAGACCCAGGGGCTGGCAAG | 1266 |
| XX | disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and | Db | 1086 | GCTGACCTTAAGTCAGACAAGCCAACTTTATATCCAGGAGACCCAGGGGCTGGCAAG | 1145 |
| XX | reproductive disorders, or disorders of the placenta. They are also | Qy | 1267 | AATTCAAAAGATGCAACCCATTCGGGGCAACAAGTCCGTTGACAAACCTTGAACCTGCG | 1326 |
| XX | useful in assessing the effects of exogenous compounds on the | Db | 1146 | AATTCAAAAGATGCAACCCATTCGGGGCAACAAGTCCGTTGACAAACCTTGAACCTGCG | 1205 |
| XX | expression of nucleic acid and amino acid sequences of CGDD. The | Qy | 1327 | AAGGAAGGCAACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGG | 1386 |
| XX | of nucleic acid and amino acid sequences of CGDD. The CGDD or its | Db | 1206 | AAGGAAGGCAACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGG | 1265 |
| XX | fragments are useful in screening compounds for effectiveness as an | Qy | | | |
| XX | agonist or antagonist of the polypeptides, or in altering the | Db | | | |
| XX | expression of the target polynucleotide and compounds that | Qy | | | |
| XX | specifically bind to or | Db | | | |
| XX | modulate the activity of the polypeptide. Microarrays | Qy | | | |
| XX | consisting of the polynucleotides of the invention are | Db | | | |
| XX | useful in monitoring or measuring | Qy | | | |
| XX | protein-protein interactions, drug-target interactions, and | Db | | | |
| XX | gene expression profiles. Sequences given in records | Qy | | | |
| XX | ACC90578-ACC90634 | Db | | | |
| XX | represent polynucleotides encoding CGDD proteins of the | Qy | | | |
| XX | invention | Db | | | |
| XX | Sequence 2020 BP; 621 A; 525 C; 529 G; 345 T; 0 U; 0 | Qy | | | |
| XX | Other; | Db | | | |
| XX | Query Match 62.9%; Score 1929.4; DB 8; Length 2020; | Qy | | | |
| XX | Best Local Similarity 96.5%; Pred. No. 0; | Db | | | |
| XX | Matches 2009; Conservative 0; Mismatches 6; Indels 66; Gaps 1; | Qy | | | |
| XX | 127 GGCAGGAGCTCCAGCAGCCAGGCACTGGGAGCAGCGACTGGAGACCCAGGACCCCT | Db | | | |
| XX | 186 | Qy | | | |
| XX | 6 GGAGCAGGCTCCACCGACGACGAGCAGTGGGAGCAGCGACTGGAGACCCAGGACCCCT | Db | | | |
| XX | 65 | Qy | | | |
| XX | 187 GTGAGGAGCAGCTCCGGGTGACACGAGGAGACTGAAGATACTCCCAAGGGGCTCAGCA | Db | | | |
| XX | 246 | Qy | | | |
| XX | 56 GTGAGGAGCAGCTCCGGGTGACACGAGGAGACTGAAGATACTCCCAAGGGGCTCAGCA | Db | | | |
| XX | 125 | Qy | | | |
| XX | 247 GGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAAGCAAGAGATGAACCA | Db | | | |
| XX | 306 | Qy | | | |

| | | | |
|----|------|---|------|
| Qy | 2064 | TCCATCTCCTCCCCAAACAGCGTCCATGTATATATTCTTCGTATGGCCAGCAATAAAT | 2123 |
| Db | 1188 | TCCAATCTCCTCCCCAAACAGCGTCCATGTATATATTCTTCGTATGGCCAGCAATAAAT | 1247 |
| Qy | 2124 | TCTSCCTAGAAAATTAAGCCGAGCTGTGTATATTAGAGGTGTATTTACGTCCTCGGT | 2183 |
| Db | 1248 | TCTSCCTAGAAAATTAAGCCGAGCTGTGTATATTAGAGGTGTATTTACGTCCTCGGT | 1307 |
| Qy | 2184 | CCAGTCTTTTCTGGCAAATAACAGTAAGAATGGTTTTAGCAGGTCACTAGTTGGGTGAGA | 2243 |
| Db | 1308 | CCAGTCTTTTCTGGCAAATAACAGTAAGAATGGTTTTAGCAGGTCACTAGTTGGGTGAGA | 1367 |
| Qy | 2244 | AGAGTCGATGATCACCAAGCAGGAAGAGGAGGGAATAGAGAAATGTGTTCCGGTTAAGTG | 2303 |
| Db | 1368 | AGAGTCGATGATCACCAAGCAGGAAGAGGAGGGAATAGAGAAATGTGTTCCGGTTAAGTG | 1427 |
| Qy | 2304 | ATGAAAAATGCAGTGGTGGCCGGCGGTGGCTCTCGCTGTAACTCTCAGCACTTTGGG | 2363 |
| Db | 1428 | ATGAAAAATGCAGTGGTGGCCGGCGGTGGCTCTCGCTGTAACTCTCAGCACTTTGGG | 1487 |
| Qy | 2364 | AGGCCGAGGCGAGGTGGATCACTGAGGTCAAGAGTTCAAGACTAGCTGCCCACAATCAT | 2423 |
| Db | 1488 | AGGCCGAGGCGAGGTGGATCACTGAGGTCAAGAGTTCAAGACTAGCTGCCCACAATCAT | 1547 |
| Qy | 2424 | GAAACCCCGTCTCTACTTAAAAATACAAAAATTAGCCAGGCAATGGTGCGCACACCTGTAG | 2483 |
| Db | 1548 | GAAACCCCGTCTCTACTTAAAAATACAAAAATTAGCCAGGCAATGGTGCGCACACCTGTAG | 1607 |
| Qy | 2484 | TCCCAGTCTACTCGGAGGCCAACCCAGCAGAACCGCTTGTATCCAGCAGGTGGAGGTTGC | 2543 |
| Db | 1608 | TCCCAGTCTACTCGGAGGCCAACCCAGCAGAACCGCTTGTATCCAGCAGGTGGAGGTTGC | 1667 |
| Qy | 2544 | AGTGAGCGCAAGTTGCACATTGCATCTCCACCTTGGCGCAGACAGCAAGATTCTATC-AA | 2602 |
| Db | 1668 | AGTGAGCGCAAGTTGCACATTGCATCTCCACCTTGGCGCAGACAGCAAGATTCTATCAAA | 1727 |
| Qy | 2603 | AAAAAAAGGCAGTGGCAAGTAAGTTATAGAGAAATGCTGCTAGAGAAATTAAGCG | 2662 |
| Db | 1728 | AAAAAAAGGCAGTGGCAAGTAAGTTATAGAGAAATGCTGCTAGAGAAATTAAGCG | 1787 |
| Qy | 2663 | TTGTAGTAAAGCGGTGCTCATCTCTAAGCTTGAAGAGGAGACAGAAATCCATTTGTT | 2722 |
| Db | 1788 | TTGTAGTAAAGCGGTGCTCATCTCTAAGCTTGAAGAGGAGACAGAAATCCATTTGTT | 1847 |
| Qy | 2723 | TAAATTACATCTCAAGAGGAGGAAACCCGGGCTGTGTGGTGGTTGCCAATTTCCCTAG | 2782 |
| Db | 1848 | TAAATTACATCTCAAGAGGAGGAAACCCGGGCTGTGTGGTGGTTGCCAATTTCCCTAG | 1907 |
| Qy | 2783 | AACCGAATGTGGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTTCAATAGGGTCC | 2842 |
| Db | 1908 | AACCGAATGTGGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTTCAATAGGGTCC | 1967 |
| Qy | 2843 | TTGTAGGTTATTGATGAGAGGAAAGATTGACCTGGGAGGCTTAAATGATTTGGGAA | 2902 |
| Db | 1968 | TTGTAGGTTATTGATGAGAGGAAAGATTGACCTGGGAGGCTTAAATGATTTGGGAA | 2027 |
| Qy | 2903 | AACAAATGCTTTTGAGGCTCAGTGACAAACCGCAAGATTACAACTTAAAAAAGAAAAA | 2962 |
| Db | 2028 | AACAAATGCTTTTGAGGCTCAGTGACAAACCGCAAGATTACAACTTAAAAAAGAAAAA | 2087 |
| Qy | 2963 | AAAAA 2967 | |
| Db | 2088 | AAAAA 2092 | |

RESULT 11
ABA17779
ID ABA17779 standard; DNA: 9889 BP.

| | |
|----|--|
| DE | Human nervous system related polynucleotide SEQ ID NO 10110. |
| XX | |
| KW | Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide; |
| KW | immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary; |
| KW | anti-parasitoid; antiskilling; antianaemic; antiarthritic; cancer; |
| KW | antihumetic; hepatotropic; cerebroprotective; antiinflammatory; |
| KW | antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal; |
| KW | antiparasitic; cardiatic; immune disorder; cardiovascular disorder; |
| KW | neurological disease; infection; neurotropic; gene therapy; vaccine; ds |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200159063-A2. |
| XX | |
| PD | 16-AUG-2001. |

PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251989P.
PR 05-DEC-2000; 2000US-0256719P.

PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX Disclosure; SEQ ID NO 10110; 1701pp + Sequence Listing; English.
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
CC (AB114678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9889 BP; 2944 A; 2085 C; 2204 G; 2656 T; 0 U; 0 Other;
Query Match 33.2%; Score 1016.4; DB 5; Length 9889;
Best Local Similarity 99.1%; Pred. No. 9.4e-246;
Matches 1033; Conservative 0; Mismatches 6; Indels 3; Gaps 1;
Qy 1929 CCTGGACCAAGCGGATGTTGGATGCTCAAGTGCACCAACAGACCCAGTATCCATCGGACC 1988
Db |||||
Qy 5062 CCAGGGACCAAGCGGATGTTGGATGCTCAAGTGCACCAACAGACCCAGTATCCATCGGACC 5121
Db |||||
Qy 1989 AGTTGGCAACCCCAAGTAACAAATCAGACGGTTCACCAAGTTCTCTGCCACCAAG 2048
Db |||||
Qy 5122 AGTTGGCAATCCCAAGTAACAAATCAGACGGTTCACCAAGTTCTCTGCCACCAAG 5181
Db |||||
Qy 2049 ATGTGTTCTCTTACTCTCCCTCCCAACAGCTCCATGTATATATTTCTCTGATG 2108
Db |||||
Qy 5182 ATGTGTTCTCTTACTCTCCCTCCCAACAGCTCCATGTATATATTTCTCTGATG 5241
Db |||||
Qy 2109 GCCAGCAAAATGAAATTTCTGCTAGAAATTAAGCCGAGCTCTTGTATATTGAGGTGATT 2168
Db |||||
Qy 5242 GCCAGCAAAATGAAATTTCTGCTAGAAATTAAGCCGAGCTCTTGTATATTGAGGTGATT 5301
Db |||||
Qy 2169 ATTTACGTCTCTGTCTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTAGCAGGTCA 2228
Db |||||
Qy 5302 ATTTACGTCTCTGTCTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTAGCAGGTCA 5361
Db |||||
Qy 2229 CCTAGTTGGGTCAAGAGTTCGATGATCACCACAGAGGAGGAGGAATAGAGGAATG 2288
Db |||||
Qy 5362 CCTAGTTGGGTCAAGAGTTCGATGATCACCACAGAGGAGGAGGAATAGAGGAATG 5421
Db |||||
Qy 2289 TGTTCCGGGTTAAGTCAATGCAAAATGGCAGTGTGGCCGGCGGTGGTGGCTCTCGCCTGTAA 2348
Db |||||
Qy 5422 TGTTCCGGGTTAAGTCAATGCAAAATGGCAGTGTGGCCGGCGGTGGTGGCTCTCGCCTGTAA 5481
Db |||||

Db 370 GTGGAGGTTGCGTGAGCGGAAGTTGACACATTGCTCCACCTGGCGGACAGCAAG 311
QY 2593 ATTCTATCAAAAAAAGGAGTGGCAAGTAAGTTATAGAGAAATGCTGTAGAA 2652
Db 310 ATTCTATCAAAAAAAGGAGTGGCAAGTAAGTTATAGAGAAATGCTGTAGAA 251
QY 2653 GAAATTAACGGTTGTAGTAACCGGTGCTCATCCTTAAGCTTGAGAGGAGAGCAAAA 2712
Db 250 GAAATTAACGGTTGTAGTAACCGGTGCTCATCCTTAAGCTTGAGAGGAGAGCAAAA 191
QY 2713 TCCATTTGTTTAAATTCACATCTCAAGGAGGAGAAACCGGGCTGTGTGGTGTGTTGCC 2772
Db 190 TCCATTTGTTTAAATTCACATCTCAAGGAGGAGAAACCGGGCTGTGTGGTGTGTTGCC 131
QY 2773 AATTTCCTAGAACGGAATGTGTGGGTATAGAAAAGGAATGAATAACGCTTGAAT 2832
Db 130 AATTTCCTAGAACGGAATGTGTGGGTATAGAAAAGGAATGAATAACGCTTGAAT 71
QY 2833 AATAGGTCCTTGTAAAGTTATTGATGAGAGGMAAAGATTGACTGGGAGGGCTTAAAT 2892
Db 70 GATAGGGTCC-TGTAAAGTTATTGATGAGAGGMAAAGATTGACTGGGAGGGCTTAAAT 13
QY 2893 GATTTGGGAAAA 2904
Db 12 GATTTGGGAAAA 1

RESULT 13

ACH89479/c
ID ACH89479 standard; DNA; 644 BP.

XX AC ACH89479;

XX DT 29-JUL-2004 (first entry)

XX DE Human genome derived single exon probe #22674.

XX KW Human; probe; ss; gene expression; single exon probe; microarray;

XX KM alternative splicing event; genomic alteration.

XX OS Homo sapiens.

XX PN US2003194704-A1.

XX PD 16-OCT-2003.

XX PF 03-APR-2002; 2002US-00029386.

XX PR 03-APR-2002; 2002US-00029386.

XX PA (PENN/) PENN S G.

XX PA (RANK/) RANK D R.

XX PA (HANZ/) HANZEL D K.

XX PI Penn SG, Rank DR, Hanzel DK;

XX DR WPI; 2004-119264/12.

XX PT New human genome-derived single exon nucleic acid probes useful for human
XX PT gene expression analysis, for identifying or characterizing alternative
XX PT splicing events, for assessing genomic alterations or as tools for
XX PT surveying tissues.

XX PS Claim 1; SEQ ID NO 22674; 80pp; English.

XX CC The invention relates to a nucleic acid probe for measuring human gene
XX CC expression, comprising any of the 27,400 fully defined nucleotide
XX CC sequences in the specification, or their complements or fragments, and
XX CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
XX CC fully defined in the specification. The probe is a single exon probe that
XX CC hybridises under high stringency conditions to a nucleic acid molecule
XX CC expressed in human cells or tissues. Also included are a spatially-
XX CC addressable set of single exon nucleic acid probes for measuring human

CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above). The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX SQ

Sequence 644 BP; 118 A; 186 C; 168 G; 172 T; 0 U; 0 Other;

Query Match 19.0%; Score 582.6; DB 12; Length 644;

Best Local Similarity 99.3%; Pred. No. 1.2e-136;

Matches 585; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 395 TCGACTTTGGGAATAAGTGTCAAGACGGATATATGCGCACCTTCTCCCGCCGAGCAACGG 454

Db 644 TCGACTTTGGGAATAAGTGTCAAGACGGATATATGCGCACCTTCTCCCGCCGAGCAACGG 585

QY 455 AGATAAGTGTCTGTTGCGGATGCCAACGGAAGAATCTTGGGAAAGAGCCAAACCCGAGG 514

Db 584 AGATAAGTGTCTGTTGCGGATGCCAACGGAAGAATCTTGGGAAAGAGCCAAACCCGAGG 525

QY 515 CACCAGCTGCTAAATCTCGTTTTTTTCTTGATGCTCTCTCGGCCCTGTACAGCAAGTACCG 574

Db 524 CACCAGCTGCTAAATCTCGTTTTTTTCTTGATGCTCTCTCGGCCCTGTACAGCAAGTACCG 465

QY 575 GAGACCAAGCCGACAGATTCTCCCTCTGATCAGTGAAGCTTGTGATGCTCAGCTCCCAATAAG 634

Db 464 GAGACCAAGCCGACAGATTCTCCCTCTGATCAGTGAAGCTTGTGATGCTCAGCTCCCAATAAG 405

QY 635 CTCCAGCGAAACAAAGACCCCAAGTGAGAGCTGGACACTTTCGGTGGCAGCTGGACCGGGGC 694

Db 404 CTCCAGCGAAACAAAGACCCCAAGTGAGAGCTGGACACTTTCGGTGGCAGCTGGACCGGGGC 345

QY 695 AGGACACAGATAAAACCCCGAGGCAACGCCCGCCCAAGCAAGGTCCTCTCTGCGCGCA 754

Db 344 AGGACACAGATAAAACCCCGAGGCAACGCCCGCCCAAGCAAGGTCCTCTCTGCGCGCA 285

QY 755 GGGATCCACGCTTCTCCACCTGAGACACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 814

Db 284 GGGATCCACGCTTCTCCACCTGAGACACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 225

QY 815 CCAAGGACTCCAGCTTTTTTTTGACAAATTTCTTCAAGCTGGACAAAGGAGGAGGAGGAGGAGG 874

Db 224 CCAAGGACTCCAGCTTTTTTTTGACAAATTTCTTCAAGCTGGACAAAGGAGGAGGAGGAGGAGG 165

QY 875 CAGGTGACAGCCAAACAGGAAGCCAAAGAGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 934

Db 164 CAGGTGACAGCCAAACAGGAAGCCAAAGAGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 105

QY 935 CTGGCTTATCAGGCGAGTCCGATGATGCTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 983

Db 104 CTGGCTTATCAGGCGAGTCCGATGATGCTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 56


```
RESULT 14
AAH50757
ID AAH50757 standard; cDNA; 488 BP.
XX
AC AAH50757;
XX
DT 23-AUG-2001 (first entry)
XX
DE Human tumour associated cDNA #86.
XX
KW Human; cancer specific gene expression; gene therapy;
KW age related differential expression; ss.
XX
OS Homo sapiens.
XX
PN W0200136685-A2.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-US031809.
XX
PR 17-NOV-1999; 99US-0166056P.
XX
PR 17-NOV-1999; 99US-0166106P.
XX
PA (NYXI-) NYXIS NEURO THERAPIES INC.
XX
PI Kroes RA, Moskal JR, Yamamoto H;
XX
DR WPI; 2001-355647/37.
XX
PT Novel nucleic acid molecules differentially expressed in brain cancers,
PT useful for ascertaining propensity of cell for malignant phenotype or
PT ascertaining suitability of anti-neoplastic drug candidate.
XX
PS Claim 28; Page 49; 82pp; English.
XX
CC The present invention provides the sequences of 184 cDNA fragments which
CC are differentially expressed in cancer cell depending on the age of the
CC patient. They can be used to diagnose and identify treatments for
CC cancers, particularly brain cancers such as haemangioblastoma, teratoma,
CC haemangioma, glioblastoma, schwannoma, osteoma and pinealoma. The present
CC sequence is a cancer-associated cDNA of the invention
XX
SQ Sequence 488 BP; 133 A; 108 C; 138 G; 109 T; 0 U; 0 Other;

Query Match 15.3%; Score 470.4; DB 4; Length 488;
Best Local Similarity 99.8%; Pred. No. 2.3e-108;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2142 CCGAGCTGTTGATATTGAGGTGATATTATTTACGTCCTGTCAGTCTTTTCGGCAA 2201
Db 15 CCGAGCTGTTGATATTGAGGTGATATTATTTACGTCCTGTCAGTCTTTTCGGCAA 74

Qy 2202 TAACAGTAAAGATGGTTAGCAGGTCACTAGTTGGGTGAGAGAGTCGATGATCACAA 2261
Db 75 TAACAGTAAAGATGGTTAGCAGGTCACTAGTTGGGTGAGAGAGTCGATGATCACAA 134

Qy 2262 GCAGGAAAGGAGGGAATAGAGGAATGTTTCGGGTTAAGTGATGAAATGGCAGTGGTG 2321
Db 135 GCAGGAAAGGAGGGAATAGAGGAATGTTTCGGGTTAAGTGATGAAATGGCAGTGGTG 194

Qy 2322 GCCGGGCTGTCGCTCTCGCTCTGTAATCTCAGACACTTTGGAGCCGAGGCGAGTGGAT 2381
Db 195 GCCGGGCTGTCGCTCTCGCTCTGTAATCTCAGACACTTTGGAGCCGAGGCGAGTGGAT 254

Qy 2382 CACCTGAGGTCAGGAGTTCAAGACTAGCTCGGCAACATCATGAAACCCCGTCTCTACTA 2441
Db 255 CACCTGAGGTCAGGAGTTCAAGACTAGCTCGGCAACATCATGAAACCCCGTCTCTACTA 314

Qy 2442 AAAATACAAAATTAGCCAGGATGGTGGGACACACCTGTGTAGTCCAGCTACTCGGGAGC 2501
|||||
```

```
Db 315 AAAATACAAAATTAGCCAGGATGGTGGGACACACCTGTGTAGTCCAGCTACTCGGGAGC 374
Qy 2502 CCAACGACGAGAACCGCTTTGTATACCCAGGAGGTGGAGTTGCAGTGCAGCCGAAAGTTGCAC 2561
|||||
Db 375 CCAACGACGAGAACCGCTTTGTATACCCAGGAGGTGGAGTTGCAGTGCAGCCGAAAGTTGCAC 434
Qy 2562 CATTGCATCTCCACCTCGGGCGACAGAGCAAGATTTCTATCAAAAAAAGGCG 2613
|||||
Db 435 CATTGCATCTCCACCTCGGGCGACAGAGCAAGATTTCTATCAAAAAAAGGCG 486

RESULT 15
ACH70719/c
ID ACH70719 standard; DNA; 576 BP.
XX
AC ACH70719;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #3914.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENNY) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119266/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 15; SEQ ID NO 3914; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
```


CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030194704
xx

SQ Sequence 576 BP; 151 A; 119 C; 136 G; 170 T; 0 U; 0 Other;

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 10.9% | Score 332.8; | DB 12; | Length 576; |
| Best Local Similarity | 99.4% | Pred. NO. 1.4e-73; | | |
| Matches 334; | Conservative 0; | Mismatches 2; | Indels 0; | Gaps 0; |

| | | | | |
|----|------|------------------------------------|-------------------------------|------|
| Qy | 1929 | CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGC | AAACAGACCCAGTATCCATCGGACC | 1988 |
| Db | 336 | CAAGGACCAAGCGGATGTTGGATGCTCAAGTGC | AAACAGACCCAGTATCCATCGGACC | 277 |
| Qy | 1989 | AGTTGGCAAAACCCAAAGTAACAAATCAGCAGG | TTCCCAACAGCTTCTCTGCCACCAAG | 2048 |
| Db | 276 | AGTTGGCAAAATCCAAAGTAACAAATCAGCAGG | TTCCCAACAGCTTCTCTGCCACCAAG | 217 |
| Qy | 2049 | ATGTGTTCTCCTTACTCCATCTCTCTCCCAAC | ACAGCTCCATGTATATATTTCTTCTGATG | 2108 |
| Db | 216 | ATGTGTTCTCCTTACTCCATCTCTCTCCCAAC | ACAGCTCCATGTATATATTTCTTCTGATG | 157 |
| Qy | 2109 | GCCAGCAAAATGAAATTCGCTAGAAATTAAGCC | CGAGCTGTTGTATATTCAGGTGTATT | 2168 |
| Db | 156 | GCCAGCAAAATGAAATTCGCTAGAAATTAAGCC | CGAGCTGTTGTATATTCAGGTGTATT | 97 |
| Qy | 2169 | ATTTACGTCCTCGTCCAGTCTTTTCTGGCAAA | ATAACAGTAAAGATGGTTTAGCAGGTCA | 2228 |
| Db | 96 | ATTTACGTCCTCGTCCAGTCTTTTCTGGCAAA | ATAACAGTAAAGATGGTTTAGCAGGTCA | 37 |
| Qy | 2229 | CCTAGTTGGGTCAGAAGATCGATGATCACCAGCA | | 2264 |
| Db | 36 | CCTAGTTGGGTCAGAAGATCGATGATCACCAGCA | | 1 |

Search completed: July 3, 2005, 10:07:13
Job time : 1575.51 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:43:42 ; Search time 483.521 Seconds
(without alignments)
10375.608 Million cell updates/sec

Title: US-08-731-499-12
Perfect score: 3066
Sequence: 1 GGAACAGCTATGACCATGA.....GGTACCCCAATTCGCCCTATA 3066
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgm2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgm2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgm2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgm2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgm2_6/ptodata/1/ina/PCTUS COMB.seq:*
6: /cgm2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|----------------------|
| 1 | 3066 | 100.0 | 3066 | 4 | US-08-892-695-12 |
| 2 | 2813 | 91.7 | 2813 | 3 | US-09-689-255C-3 |
| 3 | 2813 | 91.7 | 2813 | 4 | US-09-949-016-531 |
| 4 | 2813 | 91.7 | 2813 | 4 | US-09-949-016-4308 |
| 5 | 1018.4 | 33.2 | 130563 | 4 | US-09-949-016-12273 |
| 6 | 1018.4 | 33.2 | 131379 | 4 | US-09-949-016-16050 |
| 7 | 464.4 | 15.1 | 601 | 4 | US-09-949-016-28396 |
| 8 | 464.4 | 15.1 | 601 | 4 | US-09-949-016-153520 |
| 9 | 280 | 9.1 | 601 | 4 | US-09-949-016-153435 |
| 10 | 238 | 7.8 | 297 | 4 | US-09-513-999C-14399 |
| 11 | 236.8 | 7.7 | 72128 | 4 | US-09-949-016-16018 |
| 12 | 234 | 7.6 | 42053 | 4 | US-09-949-016-15924 |
| 13 | 232.2 | 7.6 | 137753 | 4 | US-09-949-016-17404 |
| 14 | 231.8 | 7.6 | 601 | 4 | US-09-949-016-201956 |
| 15 | 231 | 7.5 | 101011 | 4 | US-09-949-016-16933 |
| 16 | 230.8 | 7.5 | 33712 | 4 | US-09-949-016-15793 |
| 17 | 230.8 | 7.5 | 63930 | 4 | US-09-949-016-12270 |
| 18 | 230.8 | 7.5 | 109690 | 4 | US-09-949-016-13525 |
| 19 | 230.8 | 7.5 | 670689 | 4 | US-09-949-016-12505 |
| 20 | 230.8 | 7.5 | 670690 | 4 | US-09-949-016-14207 |
| 21 | 230.2 | 7.5 | 601 | 4 | US-09-949-016-119350 |
| 22 | 230.2 | 7.5 | 601 | 4 | US-09-949-016-119392 |
| 23 | 230 | 7.5 | 14345 | 4 | US-09-949-016-15449 |
| 24 | 230 | 7.5 | 39243 | 4 | US-09-949-016-12316 |
| 25 | 230 | 7.5 | 39243 | 4 | US-09-949-016-15443 |
| 26 | 229.8 | 7.5 | 152132 | 4 | US-09-949-016-13845 |
| 27 | 229.8 | 7.5 | 152145 | 4 | US-09-949-016-12371 |

| | | | | | | |
|------|-------|-----|--------|---|----------------------|--------------------|
| C 28 | 229.8 | 7.5 | 194937 | 4 | US-09-949-016-17032 | Sequence 17032, A |
| C 29 | 229.8 | 7.5 | 194937 | 4 | US-09-949-016-17033 | Sequence 17033, A |
| C 30 | 229.6 | 7.5 | 282 | 1 | US-08-133-629-8 | Sequence 8, Appli |
| C 31 | 229.6 | 7.5 | 601 | 4 | US-09-949-016-131748 | Sequence 131748, A |
| C 32 | 229.6 | 7.5 | 107085 | 4 | US-09-949-016-13157 | Sequence 13157, A |
| C 33 | 229.2 | 7.5 | 58593 | 4 | US-09-949-016-12232 | Sequence 12232, A |
| C 34 | 229.2 | 7.5 | 51403 | 4 | US-09-949-016-15057 | Sequence 15057, A |
| C 35 | 228.8 | 7.5 | 125672 | 4 | US-09-949-016-16956 | Sequence 16956, A |
| C 36 | 228.6 | 7.5 | 601 | 4 | US-09-949-016-13931 | Sequence 13931, A |
| C 37 | 228.6 | 7.5 | 601 | 4 | US-09-949-016-138463 | Sequence 138463, A |
| C 38 | 228.6 | 7.5 | 22634 | 4 | US-09-949-016-16686 | Sequence 16686, A |
| C 39 | 228.6 | 7.5 | 62804 | 3 | US-09-800-960-3 | Sequence 3, Appli |
| C 40 | 228.6 | 7.5 | 62804 | 4 | US-10-096-960-3 | Sequence 3, Appli |
| C 41 | 228.4 | 7.4 | 601 | 4 | US-09-949-016-135308 | Sequence 135308, A |
| C 42 | 228.4 | 7.4 | 26115 | 4 | US-09-949-016-13599 | Sequence 13599, A |
| C 43 | 228.4 | 7.4 | 107421 | 4 | US-09-949-016-15532 | Sequence 15532, A |
| C 44 | 228.2 | 7.4 | 97423 | 4 | US-09-949-016-12742 | Sequence 12742, A |
| C 45 | 228.2 | 7.4 | 97424 | 4 | US-09-949-016-15576 | Sequence 15576, A |

ALIGNMENTS

RESULT 1

US-08-892-695-12
; Sequence 12, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 12
; LENGTH: 3066
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:1b1
US-08-892-695-12

Query Match 100.0%; Score 3066; DB 4; Length 3066;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3066; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | GGAACAGCTATGACCATGATTAGCCAAAGCTCGAAATTAACCCCTCACTAAAGGGAACAA | 60 |
| Db | 1 | GGAACAGCTATGACCATGATTAGCCAAAGCTCGAAATTAACCCCTCACTAAAGGGAACAA | 60 |
| Qy | 61 | AAGCTGGAGTCCACCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAG | 120 |
| Db | 61 | AAGCTGGAGTCCACCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAG | 120 |
| Qy | 121 | GAATTCGGCAGCGGTCCACCGAGCCAGCCAGCGCTGGGGCAGCGACTGGAGACCCAG | 180 |
| Db | 121 | GAATTCGGCAGCGGTCCACCGAGCCAGCGCTGGGGCAGCGACTGGAGACCCAG | 180 |
| Qy | 181 | GACCTGTGCGAGGAGCAGCTCCCGGGTGCACGAGGGGACTGAAGATATCTCCACAGGGGC | 240 |
| Db | 181 | GACCTGTGCGAGGAGCAGCTCCCGGGTGCACGAGGGGACTGAAGATATCTCCACAGGGGC | 240 |

| | | | |
|----|------|---|------|
| Qy | 241 | TCAGCAGGAGCAATGGGTAAACCAAAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGAAT | 300 |
| Db | 241 | TCAGCAGGAGCAATGGGTAAACCAAAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGAAT | 300 |
| Qy | 301 | GAACCAAGCAGAGACATTACCAGGACAAACGCGTCTCTCTGTAACCGGGTTCAGTGGTG | 360 |
| Db | 301 | GAACCAAGCAGAGACATTACCAGGACAAACGCGTCTCTCTGTAACCGGGTTCAGTGGTG | 360 |
| Qy | 361 | GTGTCGACCCACACACAGTTCACGACTTAGAGGAGTGCAGTTCGGGATAAGTGTCAAGACG | 420 |
| Db | 361 | GTGTCGACCCACACACAGTTCACGACTTAGAGGAGTGCAGTTCGGGATAAGTGTCAAGACG | 420 |
| Qy | 421 | GATAAATGTGCCACTTCTTCCCCCGAGACAACCGAGATAAGTGTTCGGATGCCAAC | 480 |
| Db | 421 | GATAAATGTGCCACTTCTTCCCCCGAGACAACCGAGATAAGTGTTCGGATGCCAAC | 480 |
| Qy | 481 | GGAAAGAAATCTTTGGGAAGAGGCCAAACCCGAGGACACAGCTGCTAAATCTCGTTTTTTC | 540 |
| Db | 481 | GGAAAGAAATCTTTGGGAAGAGGCCAAACCCGAGGACACAGCTGCTAAATCTCGTTTTTTC | 540 |
| Qy | 541 | TTGATGCTCTCTCGGCCGTGTAACGAGACGTACCGGAGACCAAGCCGAGATTTCATCCCTT | 600 |
| Db | 541 | TTGATGCTCTCTCGGCCGTGTAACGAGACGTACCGGAGACCAAGCCGAGATTTCATCCCTT | 600 |
| Qy | 601 | GGATCAGTGAAGCTTTGATGTTCAGCTCCCAATAAAGCTCCAGCGAAACCAAGACCAAGTGAG | 660 |
| Db | 601 | GGATCAGTGAAGCTTTGATGTTCAGCTCCCAATAAAGCTCCAGCGAAACCAAGACCAAGTGAG | 660 |
| Qy | 661 | AGCTGGACATTTCCGGTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCAGGGCAC | 720 |
| Db | 661 | AGCTGGACATTTCCGGTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCAGGGCAC | 720 |
| Qy | 721 | GCCTCGGCCCAAGACAGGTCCTCTCTGCGCCAGGGATCCACGCTTCTCCACCTGAG | 780 |
| Db | 721 | GCCTCGGCCCAAGACAGGTCCTCTCTGCGCCAGGGATCCACGCTTCTCCACCTGAG | 780 |
| Qy | 781 | ACAGGGGAGCAGGAGGAGAGTCCCTCCAAAGCCCAAGACTCCAGCTTTTGTGACAAA | 840 |
| Db | 781 | ACAGGGGAGCAGGAGGAGAGTCCCTCCAAAGCCCAAGACTCCAGCTTTTGTGACAAA | 840 |
| Qy | 841 | TTCTTCAAGCTGGAACAAGGACAGAAAAGTGCAGGTGACAGCCAAAGGAAGCCAAAG | 900 |
| Db | 841 | TTCTTCAAGCTGGAACAAGGACAGAAAAGTGCAGGTGACAGCCAAAGGAAGCCAAAG | 900 |
| Qy | 901 | AGGCGAGACATCAAGACAAGGTGGATGAGGTTCCCTGCTTATCAGGGCAGTCCGATGAT | 960 |
| Db | 901 | AGGCGAGACATCAAGACAAGGTGGATGAGGTTCCCTGCTTATCAGGGCAGTCCGATGAT | 960 |
| Qy | 961 | GTCCCTCGAGGAAGGACATAGTTGACGCAAGGAAAGAAAGAGACAAAGAACTTGGAACT | 1020 |
| Db | 961 | GTCCCTCGAGGAAGGACATAGTTGACGCAAGGAAAGAAAGAGACAAAGAACTTGGAACT | 1020 |
| Qy | 1021 | GCAGATTGCTCTGTCTCCCTGGGGACCCAGAAAGGACTGGAGACTGCAAAAGGACGATTCACG | 1080 |
| Db | 1021 | GCAGATTGCTCTGTCTCCCTGGGGACCCAGAAAGGACTGGAGACTGCAAAAGGACGATTCACG | 1080 |
| Qy | 1081 | GCAGCAGCTATAGCAGAGAATAATAATTCCATCATGATTTCTTTAAAACTCTGGTTTCA | 1140 |
| Db | 1081 | GCAGCAGCTATAGCAGAGAATAATAATTCCATCATGATTTCTTTAAAACTCTGGTTTCA | 1140 |
| Qy | 1141 | CCTTAACAAGCTGAAACAAAAGGACCCAGAGACACGGGTCTGAAAGTCAACCCACC | 1200 |
| Db | 1141 | CCTTAACAAGCTGAAACAAAAGGACCCAGAGACACGGGTCTGAAAGTCAACCCACC | 1200 |
| Qy | 1201 | ACTTCAGCTGACCTTAAAGTCAGACAAAGCCAACTTTATATCCCAAGGAGACCAAGGGCT | 1260 |
| Db | 1201 | ACTTCAGCTGACCTTAAAGTCAGACAAAGCCAACTTTATATCCCAAGGAGACCAAGGGCT | 1260 |
| Qy | 1261 | GGCAAGAAATTCCAAAGGATGCAACCCCATTCGGGGCACAACAGTCCGTGCAACCCCTGAA | 1320 |
| Db | 1261 | GGCAAGAAATTCCAAAGGATGCAACCCCATTCGGGGCACAACAGTCCGTGCAACCCCTGAA | 1320 |

| | | | |
|------|----|---|------|
| 1321 | Qy | CCTGCGAAGGAAAGGCAACAAAGAGAAATCAGAGCCACTCTCTCGCTCTGGCCAAACTG | 1380 |
| 1321 | Db | CCTGCGAAGGAAAGGCAACCAAGGAGAAATCAGGACCCACTCTCTCGCTCTGGCCAAACTG | 1380 |
| 1381 | Qy | TTTTTGAATAAAGTTCAGTTAAAGAGGACTCAGTCCCCACAGGTGCGAGGAGAAATGTGGTG | 1440 |
| 1381 | Db | TTTTTGAATAAAGTTCAGTTAAAGAGGACTCAGTCCCCACAGGTGCGAGGAGAAATGTGGTG | 1440 |
| 1441 | Qy | TGTGAGTCCACGATAGAGATTATAAAGTCCAAAGGAAGTAGAATCAGCTTTACAAACAGTG | 1500 |
| 1441 | Db | TGTGAGTCCACGATAGAGATTATAAAGTCCAAAGGAAGTAGAATCAGCTTTACAAACAGTG | 1500 |
| 1501 | Qy | GACCTCAACGAAGGAGATGCTGCACCTGAACCCACAGAAGCGAAACTCAAAAGAGAAGAA | 1560 |
| 1501 | Db | GACCTCAACGAAGGAGATGCTGCACCTGAACCCACAGAAGCGAAACTCAAAAGAGAAGAA | 1560 |
| 1561 | Qy | AGCAAAACCAAGAACCCTCTGATGGGTTTTCTCAGACAAATGTCAGTGAAGGGGATGGA | 1620 |
| 1561 | Db | AGCAAAACCAAGAACCCTCTGATGGGTTTTCTCAGACAAATGTCAGTGAAGGGGATGGA | 1620 |
| 1621 | Qy | GGGATCACCCACTCAGAGAAATAAATGGAAAGACTCCAGCTGCCAAACATCAGACTCC | 1680 |
| 1621 | Db | GGGATCACCCACTCAGAGAAATAAATGGAAAGACTCCAGCTGCCAAACATCAGACTCC | 1680 |
| 1681 | Qy | ACAGAAAGAAGCTATCACACCGCCAGAGCTGAAACCAACAGGAGCACCAAGAGGGTAAA | 1740 |
| 1681 | Db | ACAGAAAGAAGCTATCACACCGCCAGAGCTGAAACCAACAGGAGCACCAAGAGGGTAAA | 1740 |
| 1741 | Qy | GAGGGTCTCTGAGGACAAGAAAGTCTCAGACGGCGAGATGAAACAGGAGAAGCAACAG | 1800 |
| 1741 | Db | GAGGGTCTCTGAGGACAAGAAAGTCTCAGACGGCGAGATGAAACAGGAGAAGCAACAG | 1800 |
| 1801 | Qy | CAGGAAGCCAAAGAAACAGCCAGTSCACAGAGCAGGCCACCGTGGACACGAACCTCAC | 1860 |
| 1801 | Db | CAGGAAGCCAAAGAAACAGCCAGTSCACAGAGCAGGCCACCGTGGACACGAACCTCAC | 1860 |
| 1861 | Qy | CAGAAATGGGGAACAAGCTCCAAAGAGACCTGAGAAAGCGCAGCAGTCCCTTGGGGGCTTC | 1920 |
| 1861 | Db | CAGAAATGGGGAACAAGCTCCAAAGAGACCTGAGAAAGCGCAGCAGTCCCTTGGGGGCTTC | 1920 |
| 1921 | Qy | TTTTAAAGGCTGGGACCAAAGCGGATGTTGGATGCTCAAGTGAACACAGACCCAGTATCC | 1980 |
| 1921 | Db | TTTTAAAGGCTGGGACCAAAGCGGATGTTGGATGCTCAAGTGAACACAGACCCAGTATCC | 1980 |
| 1981 | Qy | ATCGGACCAAGTTGGCAAAACCAAGTAAACAAATCAGCAGCGTTCCACACAGGTTCTCC | 2040 |
| 1981 | Db | ATCGGACCAAGTTGGCAAAACCAAGTAAACAAATCAGCAGCGTTCCACACAGGTTCTCC | 2040 |
| 2041 | Qy | CCACCAAGATGTCTTCCTTACTCCATCTCCTCCCAACACAGCTCCATGTATATATTC | 2100 |
| 2041 | Db | CCACCAAGATGTCTTCCTTACTCCATCTCCTCCCAACACAGCTCCATGTATATATTC | 2100 |
| 2101 | Qy | TTCTGATGGCCAGCAAAATTTGCTCTAGAAATTAAGCCGAGCTGTTGTAATTGA | 2160 |
| 2101 | Db | TTCTGATGGCCAGCAAAATTTGCTCTAGAAATTAAGCCGAGCTGTTGTAATTGA | 2160 |
| 2161 | Qy | GGTGTATTATTTACGTCCTGGTCCAGTCTTTTCTGGCAATAACAGTAAAGTGTTTA | 2220 |
| 2161 | Db | GGTGTATTATTTACGTCCTGGTCCAGTCTTTTCTGGCAATAACAGTAAAGTGTTTA | 2220 |
| 2221 | Qy | GCAGGTCACTAGTGGGTGAGAAAGTGCATGATCAACCAAGCAGAAAGGGAGGAATA | 2280 |
| 2221 | Db | GCAGGTCACTAGTGGGTGAGAAAGTGCATGATCAACCAAGCAGAAAGGGAGGAATA | 2280 |
| 2281 | Qy | GAGGAATGTGTTCCGGTTTAAGTGATGAAATGGCAGTGTGGCCGGCGCTGTGGCTCTC | 2340 |
| 2281 | Db | GAGGAATGTGTTCCGGTTTAAGTGATGAAATGGCAGTGTGGCCGGCGCTGTGGCTCTC | 2340 |
| 2341 | Qy | GCCTGTAACTCAGCACTTTGGGAGCCGAGCAGGTGGATCACCTGAGGTCAAGGATTC | 2400 |
| 2341 | Db | GCCTGTAACTCAGCACTTTGGGAGCCGAGCAGGTGGATCACCTGAGGTCAAGGATTC | 2400 |
| 2401 | Qy | AAGACTAGCTGGCCCAACATCATGAAACCCCGCTCTCTACTTAAATAATACAAATAATAGCCA | 2460 |


```
Db 1081 AAGTCAGACAAAGCCAACTTTACATCCAGGAGACCACAAAGGGCTGGCAAGAAATTCACAA 1140
Qy 1276 GGATGCAACCATCGGGGCACACACAGTCCGTGACAAACCCCTGAACTCGGAAGGAAGGC 1335
Db 1141 GGATGCAACCATCGGGGCACACACAGTCCGTGACAAACCCCTGAACTCGGAAGGAAGGC 1200
Qy 1336 ACCAAGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1395
Db 1201 ACCAAGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1260
Qy 1396 GTTAAAGAGGACTCAGTCCCAAGGTGCGGAGGAGAAATGTGTGTGTGAGTCAACCAAGTA 1455
Db 1261 GTTAAAGAGGACTCAGTCCCAAGGTGCGGAGGAGAAATGTGTGTGTGAGTCAACCAAGTA 1320
Qy 1456 GAGATTATAAGTCCAAAGGAAGTAGAATCAGCCTTACAACAGTGGACCTCAACGAAGGA 1515
Db 1321 GAGATTATAAGTCCAAAGGAAGTAGAATCAGCCTTACAACAGTGGACCTCAACGAAGGA 1380
Qy 1516 GATCTGCACCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAAGCAACCAAGAAC 1575
Db 1381 GATCTGCACCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAAGCAACCAAGAAC 1440
Qy 1576 TCTCTGATGGCGTTTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCACCCTCA 1635
Db 1441 TCTCTGATGGCGTTTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCACCCTCA 1500
Qy 1636 GAAGAAATAATGGGAAGACTCAGCTGCCAAACATCAGACTCCACAGAAAGACTATC 1695
Db 1501 GAAGAAATAATGGGAAGACTCAGCTGCCAAACATCAGACTCCACAGAAAGACTATC 1560
Qy 1696 ACACCGCAGAGCTGAACCAACAGGAGCACACAGAGGGTAAAGGGCTCTCGAAG 1755
Db 1561 ACACCGCAGAGCTGAACCAACAGGAGCACACAGAGGGTAAAGGGCTCTCGAAG 1620
Qy 1756 GACAAGAGTCAGCAGCGGAGATGAACAAGCAGAGAGCAACAAGCAGGAGGCCAAAGAA 1815
Db 1621 GACAAGAGTCAGCAGCGGAGATGAACAAGCAGAGAGCAACAAGCAGGAGGCCAAAGAA 1680
Qy 1816 CCAGCCAGTCACAGAGCAGGCCACCGTGACACGAACCTCAGTGCAGATGGGGACAAG 1875
Db 1681 CCAGCCAGTCACAGAGCAGGCCACCGTGACACGAACCTCAGTGCAGATGGGGACAAG 1740
Qy 1876 CTCCAAAGAGACTGTAGAGCGCAGCAGTCCCTTGGGGCTCTTTAAAGGCTCGGA 1935
Db 1741 CTCCAAAGAGACTGTAGAGCGCAGCAGTCCCTTGGGGCTCTTTAAAGGCTCGGA 1800
Qy 1936 CCAAAGCGGATGTTGGATGCTCAAGTGCAACAGACCCAGTATCCATCGGACCAAGTTGGC 1995
Db 1801 CCAAAGCGGATGTTGGATGCTCAAGTGCAACAGACCCAGTATCCATCGGACCAAGTTGGC 1860
Qy 1996 AAACCCAGTAAACAAATCAGCAGCGGTCCACAGGTTCTCTGCCACCAAGATGTT 2055
Db 1861 AAACCCAGTAAACAAATCAGCAGCGGTCCACAGGTTCTCTGCCACCAAGATGTT 1920
Qy 2056 CTCTTACTCATCTCTCTCCCAACAGCGTCCATGTATATATTTCTCTGATGCCAGCA 2115
Db 1921 CTCTTACTCATCTCTCTCCCAACAGCGTCCATGTATATATTTCTCTGATGCCAGCA 1980
Qy 2116 AATGAAATTCCTCTAGAAATTAAGCCGAGCTGTGTATATGAGGTGTATTTTACG 2175
Db 1981 AATGAAATTCCTCTAGAAATTAAGCCGAGCTGTGTATATGAGGTGTATTTTACG 2040
Qy 2176 TCTCTGTCAGTCTTTCTGGCAAAATAACAGTAAAGTGTTTAGCAGGTCACTAGTT 2235
Db 2041 TCTCTGTCAGTCTTTCTGGCAAAATAACAGTAAAGTGTTTAGCAGGTCACTAGTT 2100
Qy 2236 GGGTCAGAAGAGTCAGTATCACCAAGCAGGAGGAGGGAATAGAGGAATGTTCGG 2295
Db 2101 GGGTCAGAAGAGTCAGTATCACCAAGCAGGAGGAGGGAATAGAGGAATGTTCGG 2160
Qy 2296 GTTAAAGTATGAAATGSCAGTGTGGCCGGCGGTGGTCTCGCTGTGTAATCTCAGC 2355
```

```
Db 2161 GTTAAAGTATGAAATGCGCAGTGTGGCCGGCGGTGGTGGCTCTCGCTCTTAATCTCAGC 2220
Qy 2356 ACTTTGGAGCGCGAGCAGGTGGATCACTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 2415
Db 2221 ACTTTGGAGCGCGAGCAGGTGGATCACTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 2280
Qy 2416 AACATCATGAAACCCCGTCTCTACTAAAAATACAAAAATAGCAGGATGGTGGCACAC 2475
Db 2281 AACATCATGAAACCCCGTCTCTACTAAAAATACAAAAATAGCAGGATGGTGGCACAC 2340
Qy 2476 ACTGTAGTCCAGCTACTCGGAGGCCAAACGACGAGAACCGCTTGTACCCAGGAGGTG 2535
Db 2341 ACTGTAGTCCAGCTACTCGGAGGCCAAACGACGAGAACCGCTTGTACCCAGGAGGTG 2400
Qy 2536 GAGGTTGAGTGGAGCGGAGTGCACCATTTGCACCATTTGCACCATTTGCACCATTTGCACCAT 2595
Db 2401 GAGGTTGAGTGGAGCGGAGTGCACCATTTGCACCATTTGCACCATTTGCACCATTTGCACCAT 2460
Qy 2596 CTATCAAAAAAAGGAGCAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGTAGAGGAA 2655
Db 2461 CTATCAAAAAAAGGAGCAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGTAGAGGAA 2520
Qy 2656 TTAAGCGTTGTAGTAAACGCGTCTCATCTCTTAAGCTTGAAGAGGAGGACGAAAAATCC 2715
Db 2521 TTAAGCGTTGTAGTAAACGCGTCTCATCTCTTAAGCTTGAAGAGGAGGACGAAAAATCC 2580
Qy 2716 ATTGTGTTAAATTCACATCTCAAGGAGGAGAACCCGGCTGTGTGGGTGGTTGCCAAT 2775
Db 2581 ATTGTGTTAAATTCACATCTCAAGGAGGAGAACCCGGCTGTGTGGGTGGTTGCCAAT 2640
Qy 2776 TTCTTAGAACCGAATGTGTGGGTATAGAAAAAGGAATGAATAGCCGTGTTTTTCAAT 2835
Db 2641 TTCTTAGAACCGAATGTGTGGGTATAGAAAAAGGAATGAATAGCCGTGTTTTTCAAT 2700
Qy 2836 AGGTCCTTGTAAAGTTATTCATGAGAGGAGAAAGATTCAGTCTGGGAGGGCTTAAAAATGAT 2895
Db 2701 AGGTCCTTGTAAAGTTATTCATGAGAGGAGAAAGATTCAGTCTGGGAGGGCTTAAAAATGAT 2760
Qy 2896 TTGGGAAAAAATGCTTTTGGGCTCAGTGACAAACGCAAGGATTAACAAT 2948
Db 2761 TTGGGAAAAAATGCTTTTGGGCTCAGTGACAAACGCAAGGATTAACAAT 2813
```

RESULT 3

```
US-09-949-016-531
; Sequence 531, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 531
; LENGTH: 2813
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-531
```

```
Query Match 91.7%; Score 2813; DB 4; Length 2813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 136 CTCACCAGCAGCCAGGCACTGGGCGACGCACTGGAGACCCAGGACCCCTGTGCAGAG 195
```

Db 1 |||||CTCCACCGACGACGAGGACCTGGGCGACGACGACCTGGAGACCCAGGACCCCTGTGCGAGG 60
Qy 196 CAGCTCCGGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCAGGAGCAATG 255
Db 61 CAGCTCCGGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCAGGAGCAATG 120
Qy 256 GGTAAACCAATGAGTGTTCCTCCNAAGAGTTGAGACCAAGAGTAAGTGAACCAAGAGGACGAG 315
Db 121 GGTAAACCAATGAGTGTTCCTCCNAAGAGTTGAGACCAAGAGTAAGTGAACCAAGAGGACGAG 180
Qy 316 ACTTACCAAGGACAAACCGCTGCTGCTGAACGGGGTTCCAGTGGTGTGCTGCACCCACACA 375
Db 181 ACTTACCAAGGACAAACCGCTGCTGCTGAACGGGGTTCCAGTGGTGTGCTGCACCCACACA 240
Qy 376 GTTCAGCACTTAGAGGAAGTCCAGTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 435
Db 241 GTTCAGCACTTAGAGGAAGTCCAGTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 300
Qy 436 TCTTCCCGGAGACAAACCGGATAGTCTGCTGAACCGGGTTCGCGGATGCCAACCGGAAAGTCTTGG 495
Db 301 TCTTCCCGGAGACAAACCGGATAGTCTGCTGAACCGGGTTCGCGGATGCCAACCGGAAAGTCTTGG 360
Qy 496 AAAGAGGCCAAACCGGAGCCAGCTGCTAAATCTCGTTTTTTCTTGATGCTCTCTCGG 555
Db 361 AAAGAGGCCAAACCGGAGCCAGCTGCTAAATCTCGTTTTTTCTTGATGCTCTCTCGG 420
Qy 556 CCTGTACCAAGGACGTACCGGAGACCAAGCCGAGATTCATCCCTTTGGATCAGTGAAGCTT 615
Db 421 CCTGTACCAAGGACGTACCGGAGACCAAGCCGAGATTCATCCCTTTGGATCAGTGAAGCTT 480
Qy 616 GATGTACGCTCCATTAAGCTCCAGCGAAACAAAGACCCCAAGTGAGAGCTGACACTTCCG 675
Db 481 GATGTACGCTCCATTAAGCTCCAGCGAAACAAAGACCCCAAGTGAGAGCTGACACTTCCG 540
Qy 676 GTGCGAGCTGACCGGGGACGACACAGATAAACCCAGGCGACGCCCCGGGCCAAGAC 735
Db 541 GTGCGAGCTGACCGGGGACGACACAGATAAACCCAGGCGACGCCCCGGGCCAAGAC 600
Qy 736 AAGGTCTCTCTGCGCGCAGGGATCCCAAGCTTCTCCACCTTGACACAGGGGGAGCAGGA 795
Db 601 AAGGTCTCTCTGCGCGCAGGGATCCCAAGCTTCTCCACCTTGACACAGGGGGAGCAGGA 660
Qy 796 GGAGAGCTCCCTCCAAAGCCCAAGGACTCCAGCTTTTGTGACAAATCTTCAAGCTGAC 855
Db 661 GGAGAGCTCCCTCCAAAGCCCAAGGACTCCAGCTTTTGTGACAAATCTTCAAGCTGAC 720
Qy 856 AAGGACAGGAAAGGTGCGGAGGTGACAGCCAAAGAGGACCAAGAGGCGAGAGCATCAA 915
Db 721 AAGGACAGGAAAGGTGCGGAGGTGACAGCCAAAGAGGACCAAGAGGCGAGAGCATCAA 780
Qy 916 GACAAGGTGGATGAGTTTCTGGCTTATCAGGCGAGTCCGATGATGTCCTGCGAGGGAAG 975
Db 781 GACAAGGTGGATGAGTTTCTGGCTTATCAGGCGAGTCCGATGATGTCCTGCGAGGGAAG 840
Qy 976 GACATAGTTGACGCAAGGAAAGAGGACAAAGAACTTGGAACTGCGGATGCTCTGTC 1035
Db 841 GACATAGTTGACGCAAGGAAAGAGGACAAAGAACTTGGAACTGCGGATGCTCTGTC 900
Qy 1036 CTTGGGACCCAGAGGACTGGAGCTGCAAGGACGATTCAGGCGAGCAGCTATAGCA 1095
Db 901 CTTGGGACCCAGAGGACTGGAGCTGCAAGGACGATTCAGGCGAGCAGCTATAGCA 960
Qy 1096 GAGATAATAATTCATCATGAGTTTCTTAAACTCTGTTTACCTTAAACAAAGCTCAA 1155
Db 961 GAGATAATAATTCATCATGAGTTTCTTAAACTCTGTTTACCTTAAACAAAGCTCAA 1020
Qy 1156 ACAAAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACCTTACGCTGACCTT 1215
Db 1021 ACAAAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACCTTACGCTGACCTT 1080
Qy 1216 AAGTCAGACAAAGCCAACTTTATCCAGGAGACCCAGGGGCTGGCAAGATTCGAAA 1275
|||||

Db 1081 AAGTCAGACAAAGCCAACTTTACATCCAGGAGACCCAGGGCTGGCAAGATTTCCAAA 1140
Qy 1276 GGATGCAACCAATCGGGGACACACAGTTCCTGTGACAACTTCTGAACTCTGCGAAGGAAGGC 1335
Db 1141 GGATGCAACCAATCGGGGACACACAGTTCCTGTGACAACTTCTGAACTCTGCGAAGGAAGGC 1200
Qy 1336 ACCAAGGAGAAATCAGAGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAGTCA 1395
Db 1201 ACCAAGGAGAAATCAGAGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAGTCA 1260
Qy 1396 GTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGGAGAAATGTGGTGTGTGAGTCAACCACTA 1455
Db 1261 GTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGGAGAAATGTGGTGTGTGAGTCAACCACTA 1320
Qy 1456 GAGATTATAAGTCCAAAGGAAGTGAATCAGCCTTCAAAACAGTGGACCTCAACGAAGGA 1515
Db 1321 GAGATTATAAGTCCAAAGGAAGTGAATCAGCCTTCAAAACAGTGGACCTCAACGAAGGA 1380
Qy 1516 GATCTCACCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAAGCAACCAACCAAGAAC 1575
Db 1381 GATCTGACCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAAGCAACCAACCAAGAAC 1440
Qy 1576 TCTCTGATGGGCTTTCTCAGACAAATGTCACTGAAGGGGATGGAGGGATCACCCACTCA 1635
Db 1441 TCTCTGATGGGCTTTCTCAGACAAATGTCACTGAAGGGGATGGAGGGATCACCCACTCA 1500
Qy 1636 GAAGAAATAATGGGAAGACTCCAGCTGCAAAACATCAGACTCCACAGAAAAGACTATC 1695
Db 1501 GAAGAAATAATGGGAAGACTCCAGCTGCAAAACATCAGACTCCACAGAAAAGACTATC 1560
Qy 1696 ACACGGCAGAGCTGACCAACAGGAGGACCAAGAGGGTAAAGGGCTCTCTCGAAG 1755
Db 1561 ACACGGCAGAGCTGACCAACAGGAGGACCAAGAGGGTAAAGGGCTCTCTCGAAG 1620
Qy 1756 GACAAGAAAGTCAGCAGCGGAGATGAAACAGCAGAAAGAGCAACAGCAGGAAAGCCAAAGAA 1815
Db 1621 GACAAGAAAGTCAGCAGCGGAGATGAAACAGCAGAAAGAGCAACAGCAGGAAAGCCAAAGAA 1680
Qy 1816 CCAGCCAGTGCAACAGAGCAGGCGCCAGCTGCAACAGCAGAACTCACTGCAAGATGGGACAAAG 1875
Db 1681 CCAGCCAGTGCAACAGAGCAGGCGCCAGCTGCAACAGCAGAACTCACTGCAAGATGGGACAAAG 1740
Qy 1876 CTCCAAAAGAGACCTGAGAGCGGACAGCTCTTGGGGCTCTTTAAAGGCTGGA 1935
Db 1741 CTCCAAAAGAGACCTGAGAGCGGACAGCTCTTGGGGCTCTTTAAAGGCTGGA 1800
Qy 1936 CCAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGACCAAGTTGGC 1995
Db 1801 CCAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGACCAAGTTGGC 1860
Qy 1996 AAACCCAGTAAACAAATCAGCAGGTTCCACAGGTTCTCTGCGCAACCAAGATGTGT 2055
Db 1861 AAACCCAGTAAACAAATCAGCAGGTTCCACAGGTTCTCTCTGCGCAACCAAGATGTGT 1920
Qy 2056 CTCTTACTCCATCTCTCTCCCAACAGGCTCCATGTATATATTTCTTCTGATGGCCAGA 2115
Db 1921 CTCTTACTCCATCTCTCTCCCAACAGGCTCCATGTATATATTTCTTCTGATGGCCAGA 1980
Qy 2116 AATCAAAATTCGCTTAGAAATTAAGCCCGAGCTGTTGTATATTTAGAGTGTATTTATTACG 2175
Db 1981 AATCAAAATTCGCTTAGAAATTAAGCCCGAGCTGTTGTATATTTAGAGTGTATTTATTACG 2040
Qy 2176 TCTCTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTTAGCAGGTCACCTAGTT 2235
Db 2041 TCTCTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTTAGCAGGTCACCTAGTT 2100
Qy 2236 GGGTCAGAGAGTGCATGATCAACAGAGGAAAGGGAGGGAATAGAGGAATGTGTTCCG 2295
Db 2101 GGGTCAGAGAGTGCATGATCAACAGAGGAAAGGGAGGGAATAGAGGAATGTGTTCCG 2160
Qy 2296 GTTAAAGTCAAAATGCGAGTGGTGGCGGGCTGTTGGCTCTCGCTGTAACTCAGC 2355
Db 2161 GTTAAAGTCAAAATGCGAGTGGTGGCGGGCTGTTGGCTCTCGCTGTAACTCAGC 2220
|||||

```
QY 2356 ACTTTGGAGGCCGAGCAGGTGGATCACCTGAGGTGAGGATTTCAAGACTAGCCTGGCC 2415
Db |||||
QY 2221 ACTTTGGAGGCCGAGCAGGTGGATCACCTGAGGTGAGGATTTCAAGACTAGCCTGGCC 2280
Db |||||
QY 2416 AACATCATGAACCCCGTCTCTACTATAAATAACAAAAATAGCCAGGCATGGTGGCACAC 2475
Db |||||
QY 2281 AACATCATGAACCCCGTCTCTACTATAAATAACAAAAATAGCCAGGCATGGTGGCACAC 2340
Db |||||
QY 2476 ACCTGTAGTCCAGTACTCGGAGCCCAACGACGAGAACCGCTTGTACCCAGAGGTG 2535
Db |||||
QY 2341 ACCTGTAGTCCAGTACTCGGAGCCCAACGACGAGAACCGCTTGTACCCAGAGGTG 2400
Db |||||
QY 2536 GAGGTTGCAGTGAGCCGAAGTTGCAACATTGCACTCCACCTCGGCGCACGAGCAAGATT 2595
Db |||||
QY 2401 GAGGTTGCAGTGAGCCGAAGTTGCAACATTGCACTCCACCTCGGCGCACGAGCAAGATT 2460
Db |||||
QY 2596 CTATCAAAAAAAGGAGGAGTGGCAAGTAAGTATAGAGAAATGCTGCTGAAGGAA 2655
Db |||||
QY 2461 CTATCAAAAAAAGGAGGAGTGGCAAGTAAAGTTATAGAGAAATGCTGCTGAAGGAA 2520
Db |||||
QY 2656 TTAAGCGTTGTAGTAAACGGTGTCTATCTTAAGCTTGAAGAGGAGACGAAATCC 2715
Db |||||
QY 2521 TTAAGCGTTGTAGTAAACGGTGTCTATCTTAAGCTTGAAGAGGAGACGAAATCC 2580
Db |||||
QY 2716 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGTGTGTGGGTGTTGCCAAT 2775
Db |||||
QY 2581 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGTGTGTGGGTGTTGCCAAT 2640
Db |||||
QY 2776 TTCTAGAACCGAATGTGTGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAAAT 2835
Db |||||
QY 2641 TTCTAGAACCGAATGTGTGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAAAT 2700
Db |||||
QY 2836 AGGTCCTTGTAGTAAATGATGAGAGGAGAAAGATTGACTGGGAGGCTTAAATGAT 2895
Db |||||
QY 2701 AGGTCCTTGTAGTAAATGATGAGAGGAGAAAGATTGACTGGGAGGCTTAAATGAT 2760
Db |||||
QY 2896 TTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACCGCAAGATTACAACTT 2948
Db |||||
QY 2761 TTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACCGCAAGATTACAACTT 2813
Db |||||

RESULT 4
US-09-949-016-4308
; Sequence 4308, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4308
; LENGTH: 3629
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4308

Query Match 91.7%; Score 2813; DB 4; Length 3629;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2827; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 136 CTCACCGACAGCCAGGCACTGGGCGAGCAGCAGCTGGAGACCCAGGACCCCTGTGCAGGAG 195
Db |||||
```

```
Db 1 CTCACCGACAGCCAGGCACTGGGCGAGCAGCAGCTGGAGACCCAGGACCCCTGTGCAGGAG 60
QY 196 CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCACACAGGGGCTCAGCAGGAGCAATG 255
Db |||||
QY 61 CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCACACAGGGGCTCAGCAGGAGCAATG 120
Db |||||
QY 256 GGTAAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGAAATGAACAGAGACGAG 315
Db |||||
QY 121 GGTAAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGAAATGAACAGAGACGAG 180
Db |||||
QY 316 ACTTACGAGACAAACCGTCTCTGTAACGGGTTCCAGTGTGTGTGTCGACCCACACA 375
Db |||||
QY 181 ACTTACGAGACAAACCGTCTCTGTAACGGGTTCCAGTGTGTGTGTCGACCCACACA 240
Db |||||
QY 376 GTTCAGCACTTAGAGGAAGTTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 435
Db |||||
QY 241 GTTCAGCACTTAGAGGAAGTTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 300
Db |||||
QY 436 TCTTCCCGGAGACAAACCGGAGATAAGTGTGTGGGATGCCAACGGAAAGAAATCTTGGG 495
Db |||||
QY 301 TCTTCCCGGAGACAAACCGGAGATAAGTGTGTGGGATGCCAACGGAAAGAAATCTTGGG 360
Db |||||
QY 496 AAAGAGCCCAAAACCGGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTT 615
Db |||||
QY 421 CTTGTACAGAGACGTAACCGGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTT 480
Db |||||
QY 616 GATGTACAGCTTCCAATAAGCTCCAGCGAAACAAAGACCCAAAGTGAAGCTGAGACACTTCCG 675
Db |||||
QY 481 GATGTACAGCTTCCAATAAGCTCCAGCGAAACAAAGACCCAAAGTGAAGCTGAGACACTTCCG 540
Db |||||
QY 676 GTGCAGCTGGAACCGGCGAGGACACAGATAAACCAGGCGACGCCCGGCCCAAGAC 735
Db |||||
QY 541 GTGCAGCTGGAACCGGCGAGGACACAGATAAACCAGGCGACGCCCGGCCCAAGAC 600
Db |||||
QY 736 AAGTGTCTCTCTGCGCCAGGATCCACGCTTCTCCACCTGAGACAGGGGAGCAGGA 795
Db |||||
QY 601 AAGTGTCTCTCTGCGCCAGGATCCACGCTTCTCCACCTGAGACAGGGGAGCAGGA 660
Db |||||
QY 796 GGAGAACTCCCTCCAAAGCCAAAGGACTCCAGCTTTTGTGACAAATTTCTCAAGCTGGAC 855
Db |||||
QY 661 GGAGAACTCCCTCCAAAGCCAAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGAC 720
Db |||||
QY 856 AAGGACAGGAAAGGTGCCAGGTGACAGCCACAGGAAGCCAGAGGCGAGACATCAA 915
Db |||||
QY 721 AAGGACAGGAAAGGTGCCAGGTGACAGCCACAGGAAGCCAGAGGCGAGACATCAA 780
Db |||||
QY 916 GACAAGTGTGATGAGGTTCTCTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCAAGGAAAG 975
Db |||||
QY 781 GACAAGTGTGATGAGGTTCTCTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCAAGGAAAG 840
Db |||||
QY 976 GACATAGTTGACGCAAGGAAAAAGAGGACAAAGAACTTGGAACTGCGGATGCTGTCTC 1035
Db |||||
QY 841 GACATAGTTGACGCAAGGAAAAAGAGGACAAAGAACTTGGAACTGCGGATGCTGTCTC 900
Db |||||
QY 1036 CTTGGGACCCAGAGGACTGGAGACTGCAAGAGACGATTTCCAGGCGAGCAGCTATAGCA 1095
Db |||||
QY 901 CTTGGGACCCAGAGGACTGGAGACTGCAAGAGACGATTTCCAGGCGAGCAGCTATAGCA 960
Db |||||
QY 1096 GAGAAATAATTCATCATGAGTTTCTTTTAAAACTCTGGTTTCACTTAAACAAAGCTGAA 1155
Db |||||
QY 961 GAGAAATAATTCATCATGAGTTTCTTTTAAAACTCTGGTTTCACTTAAACAAAGCTGAA 1020
Db |||||
QY 1156 AAAAAAGGACCCAGAGGACACGGGTGCTGAAAAAGTCAACCACTTCAAGCTGACCTT 1215
Db |||||
QY 1021 AAAAAAGGACCCAGAGGACACGGGTGCTGAAAAAGTCAACCACTTCAAGCTGACCTT 1080
Db |||||
QY 1216 AAGTCAGACAAAGCAACTTTACATCCAGAGACCCAGGGCTGGCAAGAAATTCCAA 1275
Db |||||
QY 1081 AAGTCAGACAAAGCAACTTTTACATCCAGAGACCCAGGGCTGGCAAGAAATTCCAA 1140
Db |||||
```


| | | | |
|----|--------|---|--------|
| Qy | 1929 | CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCAGTATCCATCGGACC | 1988 |
| Db | 127543 | CCAGGACCAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCAGTATCCATCGGACC | 127602 |
| Qy | 1989 | AGTTGGCAAAACCCAAAGTAAACAAATCAGCAGCGTTCCCAACAGGTTCTCTCGCCACCAAG | 2048 |
| Db | 127603 | AGTTGGCAAAACCCAAAGTAAACAAATCAGCAGCGTTCCCAACAGGTTCTCTCGCCACCAAG | 127662 |
| Qy | 2049 | ATGTGTTCTCCTTACTCCATCTCTCCCAACAGCGTCCATGTATATATCTTCTGATG | 2108 |
| Db | 127663 | ATGTGTTCTCCTTACTCCATCTCTCCCAACAGCGTCCATGTATATATCTTCTGATG | 127722 |
| Qy | 2109 | GCACGAAATGAAATCTCGCTCAGAAATTAAGCCCGAGCTGTTCTATATATTGAGTGATT | 2168 |
| Db | 127723 | GCACGAAATGAAATCTCGCTCAGAAATTAAGCCCGAGCTGTTCTATATATTGAGTGATT | 127782 |
| Qy | 2169 | ATTTACGCTCTGCTCGAGTCTTTTCTGGCAAAATACAGTAAAGATGGTTTAGAGTCA | 2228 |
| Db | 127783 | ATTTACGCTCTGCTCGAGTCTTTTCTGGCAAAATACAGTAAAGATGGTTTAGAGTCA | 127842 |
| Qy | 2229 | CCTAGTTGGTTCAGAAAGTCGATCACCACAGGAAAGGAGGGAATAGGGAATG | 2288 |
| Db | 127843 | CCTAGTTGGTTCAGAAAGTCGATCACCACAGGAAAGGAGGGAATAGGGAATG | 127902 |
| Qy | 2289 | TGTTCCGGGTTAAGTGATGAAATGGCGGCGTGGTGGCTCTCGCTCTGTAA | 2348 |
| Db | 127903 | TGTTCCGGGTTAAGTGATGAAATGGCGGCGTGGTGGCTCTCGCTCTGTAA | 127962 |
| Qy | 2349 | TCTCAGCACTTTGGGAGGCGAGCAGGTGGATCACCTGAGGTCAGAGGTTCAAGACTAG | 2408 |
| Db | 127963 | TCTCAGCACTTTGGGAGGCGAGCAGGTGGATCACCTGAGGTCAGAGGTTCAAGACTAG | 128022 |
| Qy | 2409 | CCTGGCCAAACATCATGAAACCCCGTCTCTACTTAAATAACAAAATTTAGCCAGCATGGT | 2468 |
| Db | 128023 | CCTGGCCAAACATCATGAAACCCCGTCTCTACTTAAATAACAAAATTTAGCCAGCATGGT | 128082 |
| Qy | 2469 | GGCACACACCTGTAGTCCAGCTACTCGGGAGGCCAACGACGAGAACCGCTTGACCCA | 2528 |
| Db | 128083 | GGCACACACCTGTAGTCCAGCTACTCGGGAGGCCAACGACGAGAACCGCTTGACCCA | 128142 |
| Qy | 2529 | GGAGGTGGAGGTTGCAGTAGCCGAAAGTTGCACCATTTGCACTCCACCCTGGGCGACAG | 2588 |
| Db | 128143 | GGAGGTGGAGGTTGCAGTAGCCGAAAGTTGCACCATTTGCACTCCACCCTGGGCGACAG | 128202 |
| Qy | 2589 | CAAGATTTCTATC-NAAAAAAAAAAGGCGAGTGGCAAGTAAATGATGAAATGCTGCT | 2647 |
| Db | 128203 | CAAGATTTCTATC-NAAAAAAAAAAGGCGAGTGGCAAGTAAATGATGAAATGCTGCT | 128262 |
| Qy | 2648 | AGAGGAATTAAGCGTTGTAGTAAACCGTCTCATCTCTAAGCTTGAAGAGGGAGAC | 2707 |
| Db | 128263 | AGAGGAATTAAGCGTTGTAGTAAACCGTCTCATCTCTAAGCTTGAAGAGGGAGAC | 128322 |
| Qy | 2708 | GAAATCCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGG | 2767 |
| Db | 128323 | GAAATCCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGG | 128382 |
| Qy | 2768 | TTGCCAATTTCTAGAACGGAATGTGGGGTATAGAAAGGAATGAAATGAGGTTGTT | 2827 |
| Db | 128383 | TTGCCAATTTCTAGAACGGAATGTGGGGTATAGAAAGGAATGAAATGAGGTTGTT | 128442 |
| Qy | 2828 | TTTCAAAATAGGGTCTTGTAAAGTTATTGATCAGAGGGGAAAGATTGACTGGGAGGGCTT | 2887 |
| Db | 128443 | TTTCAAAATAGGGTCTTGTAAAGTTATTGATCAGAGGGGAAAGATTGACTGGGAGGGCTT | 128502 |
| Qy | 2888 | AAAAATGATTTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACGGCAAAAGATTACAACT | 2947 |
| Db | 128503 | AAAAATGATTTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACGGCAAAAGATTACAACT | 128562 |
| Qy | 2948 | TAATAAAAAAAAAAAAAA 2967 | |
| Db | 128563 | TAATAAAAAAAAAAAAAA 128582 | |

| | | | |
|---|--------|---|--------|
| RESULT 6 | | | |
| US-09-949-016-16050 | | | |
| ; Sequence 16050, Application US/09949016 | | | |
| ; Patent No. 6812339 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: VENTER, J. Craig et al. | | | |
| ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED | | | |
| ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF | | | |
| ; FILE REFERENCE: CL001307 | | | |
| ; CURRENT APPLICATION NUMBER: US/09/949,016 | | | |
| ; CURRENT FILING DATE: 2000-04-14 | | | |
| ; PRIOR APPLICATION NUMBER: 60/241,755 | | | |
| ; PRIOR FILING DATE: 2000-10-20 | | | |
| ; PRIOR APPLICATION NUMBER: 60/237,768 | | | |
| ; PRIOR FILING DATE: 2000-10-03 | | | |
| ; PRIOR APPLICATION NUMBER: 60/231,498 | | | |
| ; PRIOR FILING DATE: 2000-09-08 | | | |
| ; NUMBER OF SEQ ID NOS: 207012 | | | |
| ; SOFTWARE: FastSeq for Windows Version 4.0 | | | |
| ; SEQ ID NO 16050 | | | |
| ; LENGTH: 131379 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Human | | | |
| US-09-949-016-16050 | | | |
| Query Match | | | |
| Best Local Similarity 33.2%; Score 1018.4; DB 4; Length 131379; | | | |
| Matches 1033; Conservative 0; Mismatches 6; Indels 1; Gaps 1; | | | |
| Qy | 1929 | CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCAGTATCCATCGGACC | 1988 |
| Db | 127543 | CCAGGACCAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCAGTATCCATCGGACC | 127602 |
| Qy | 1989 | AGTTGGCAAAACCCAAAGTAAACAAATCAGCAGCGTTCCCAACAGGTTCTCTCGCCACCAAG | 2048 |
| Db | 127603 | AGTTGGCAAAACCCAAAGTAAACAAATCAGCAGCGTTCCCAACAGGTTCTCTCGCCACCAAG | 127662 |
| Qy | 2049 | ATGTGTTCTCCTTACTCCATCTCTCCCAACAGCGTCCATGTATATATCTTCTGATG | 2108 |
| Db | 127663 | ATGTGTTCTCCTTACTCCATCTCTCCCAACAGCGTCCATGTATATATCTTCTGATG | 127722 |
| Qy | 2109 | GCACGAAATGAAATCTCGCTCAGAAATTAAGCCCGAGCTGTTCTATATATTGAGTGATT | 2168 |
| Db | 127723 | GCACGAAATGAAATCTCGCTCAGAAATTAAGCCCGAGCTGTTCTATATATTGAGTGATT | 127782 |
| Qy | 2169 | ATTTACGCTCTGCTCGAGTCTTTTCTGGCAAAATACAGTAAAGATGGTTTAGAGTCA | 2228 |
| Db | 127783 | ATTTACGCTCTGCTCGAGTCTTTTCTGGCAAAATACAGTAAAGATGGTTTAGAGTCA | 127842 |
| Qy | 2229 | CCTAGTTGGTTCAGAAAGTCGATCACCACAGGAAAGGAGGGAATAGGGAATG | 2288 |
| Db | 127843 | CCTAGTTGGTTCAGAAAGTCGATCACCACAGGAAAGGAGGGAATAGGGAATG | 127902 |
| Qy | 2289 | TGTTCCGGGTTAAGTGATGAAATGGCGGCGTGGTGGCTCTCGCTCTGTAA | 2348 |
| Db | 127903 | TGTTCCGGGTTAAGTGATGAAATGGCGGCGTGGTGGCTCTCGCTCTGTAA | 127962 |
| Qy | 2349 | TCTCAGCACTTTGGGAGGCGAGCAGGTGGATCACCTGAGGTCAGAGGTTCAAGACTAG | 2408 |
| Db | 127963 | TCTCAGCACTTTGGGAGGCGAGCAGGTGGATCACCTGAGGTCAGAGGTTCAAGACTAG | 128022 |
| Qy | 2409 | CCTGGCCAAACATCATGAAACCCCGTCTCTACTTAAATAACAAAATTTAGCCAGCATGGT | 2468 |
| Db | 128023 | CCTGGCCAAACATCATGAAACCCCGTCTCTACTTAAATAACAAAATTTAGCCAGCATGGT | 128082 |
| Qy | 2469 | GGCACACACCTGTAGTCCAGCTACTCGGGAGGCCAACGACGAGAACCGCTTGACCCA | 2528 |
| Db | 128083 | GGCACACACCTGTAGTCCAGCTACTCGGGAGGCCAACGACGAGAACCGCTTGACCCA | 128142 |
| Qy | 2529 | GGAGGTGGAGGTTGCAGTAGCCGAAAGTTGCACCATTTGCACTCCACCCTGGGCGACAG | 2588 |
| Db | 128143 | GGAGGTGGAGGTTGCAGTAGCCGAAAGTTGCACCATTTGCACTCCACCCTGGGCGACAG | 128202 |
| Qy | 2589 | CAAGATTTCTATC-NAAAAAAAAAAGGCGAGTGGCAAGTAAATGATGAAATGCTGCT | 2647 |
| Db | 128203 | CAAGATTTCTATC-NAAAAAAAAAAGGCGAGTGGCAAGTAAATGATGAAATGCTGCT | 128262 |
| Qy | 2648 | AGAGGAATTAAGCGTTGTAGTAAACCGTCTCATCTCTAAGCTTGAAGAGGGAGAC | 2707 |
| Db | 128263 | AGAGGAATTAAGCGTTGTAGTAAACCGTCTCATCTCTAAGCTTGAAGAGGGAGAC | 128322 |
| Qy | 2708 | GAAATCCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGG | 2767 |
| Db | 128323 | GAAATCCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGG | 128382 |
| Qy | 2768 | TTGCCAATTTCTAGAACGGAATGTGGGGTATAGAAAGGAATGAAATGAGGTTGTT | 2827 |
| Db | 128383 | TTGCCAATTTCTAGAACGGAATGTGGGGTATAGAAAGGAATGAAATGAGGTTGTT | 128442 |
| Qy | 2828 | TTTCAAAATAGGGTCTTGTAAAGTTATTGATCAGAGGGGAAAGATTGACTGGGAGGGCTT | 2887 |
| Db | 128443 | TTTCAAAATAGGGTCTTGTAAAGTTATTGATCAGAGGGGAAAGATTGACTGGGAGGGCTT | 128502 |
| Qy | 2888 | AAAAATGATTTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACGGCAAAAGATTACAACT | 2947 |
| Db | 128503 | AAAAATGATTTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACGGCAAAAGATTACAACT | 128562 |
| Qy | 2948 | TAATAAAAAAAAAAAAAA 2967 | |
| Db | 128563 | TAATAAAAAAAAAAAAAA 128582 | |

| | | | |
|----|--------|---|--------|
| Qy | 2589 | CAAGATTCTATC - AAAAAAAAAAGGCAGTGGCAAGTAAGTTATAGAAGAGAATCTGCT | 2647 |
| Db | 128203 | CAAGATTCTATCAAAAAAAAAAGGCAGTGGCAAGTAAGTTATAGAAGAGAATCTGCT | 128262 |
| Qy | 2648 | AGAAGGAATTAAAGCGTTCTAGTAAACGGCTGCTCATCTCTAAGCTTGAAGAAGGGAGAC | 2707 |
| Db | 128263 | AGAAGGAATTAAAGCGTTCTAGTAAATGCTGCTTATCTCTAAGCTTGAAGAAGGGAGAC | 128322 |
| Qy | 2708 | GAAATCCATTTGTTTAAATTCACATCTCAAGAGGGAGAACCCGGGCTGTTCGGTGG | 2767 |
| Db | 128323 | GAAATCCATTTGTTTAAATTCACATCTCAAGAGGGAGAACCCGGGCTGTTCGGTGG | 128382 |
| Qy | 2768 | TTGGCAATTTCTTAGACCGGAATGTGCGGTATAGAAAAAGCGAATGAATAGCGTTGTT | 2827 |
| Db | 128383 | TTGGCAATTTCTTAGACCGGAATGTGCGGTATAGAAAAAGCGAATGAATAGCGTTGTT | 128442 |
| Qy | 2828 | TTTCAAAATAGGCTCCTTGTAAGTTATTGATGAGAGGGAAAAGATTGACTGGGGAGGGCTT | 2887 |
| Db | 128443 | TTTCAAAATAGGCTCCTTGTAAGTTATTGATGAGAGGGAAAAGATTGACTGGGGAGGGCTT | 128502 |
| Qy | 2888 | AAAATGATTTGGGAAAACAATTGCTTTTGGCTCAGTGACAA CGGCAAAAGATTACAAC | 2947 |
| Db | 128503 | AAAATGATTTGGGAAAACAATTGCTTTTGGCTCAGTGACAA CGGCAAAAGATTACAAC | 128562 |
| Qy | 2948 | TAAAAAATAAAAAAAAAAAAAA 2967 | |
| Db | 128563 | TAAAAAATAAAAAAAAAAAAAA 128582 | |

RESULT. T 7

```

US-09-949-016-28396/c
; Sequence 28396, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28396
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-28396

```

RESULT 8

```

US-09-949-016-153520/c
; Sequence 153520, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153520
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-153520

```

| | | | | | | | |
|----|------|--------------|--------------|------------------|----------------------|-----------------------|------|
| Db | 288 | GCAGCAAAATGA | AATTCGCTAGAA | TAATTAAGCCGAGCTG | TGTGTAATTAATGAGGTGTA | TT | 229 |
| Qy | 2169 | ATTTACGTC | CTCGGTC | CCAGTCTTTTCT | CTGGCAAAATAACAGTA | AGATGGTTTAGCAGGTCA | 2228 |
| Db | 228 | ATTTACGTC | CTCGGTC | CCAGTCTTTTCT | CTGGCAAAATAACAGT | AAAGATGGTTTAGCAGGTCA | 169 |
| Qy | 2229 | CCTAGTTGGG | TCAGAGAGT | CGATGATCA | CCAAGCAGGAA | AAGGAGGGAATAGAGGAATG | 2288 |
| Db | 168 | CCTAGTTGGG | TCAGAGAGT | CGATGATCA | CACACAGCAGG | AAAGGAGGGAATAGAGGAATG | 109 |
| Qy | 2289 | TGTTTGGG | TTAAGTGA | TGAAATGGCAGT | CGTGGCGGGCGTGGT | GCTCTCGCTCTGTA | 2348 |
| Db | 108 | TGTTTGGG | TTAAGTGA | TGAAATGGCAGT | CGTGGCGGGCGTGGT | GCTCTCGCTCTGTA | 49 |
| Qy | 2349 | TCTCAGCA | CTTTGGGAG | CCGAGGCAGG | TGGATCA | CCTGAGGTCAGGA | 2396 |
| Db | 48 | TCTCAGCA | CTTTGGGAG | CCGAGGCAGG | TGGATCA | CCTGAGGTCAGGA | 1 |

Qy 2109 GCCAGCAAATGAAATTCCTAGAAA'TTAAGCCGAGCTGTTGTATATTGAGGTATT 2168

```
Db 108 TGTTCGGGTTAAAGTATCAAAATGGCAGTGTGGCCGGGCTGTGGCTCTCGCCTGTAA 49
Qy 2349 TCTCAGCACTTTGGGAGCGGAGCGAGGTGGATCACCTGAGGTCAGGA 2396
Db 48 TCTCAGCACTTTGGGAGCGGAGCGAGGTGGATCACCTGAGGTCAGGA 1

RESULT 9
US-09-949-016-153435/c
; Sequence 153435, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153435
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-153435

Query Match 9.1%; Score 280; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 4.9e-64;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGGAGACAAC 452
Db 280 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGGAGACAAC 221

Qy 453 GGAGATAAGTCTGTTGGGATGCCAAGCGAAGAAATCTTGGGAAGAGGCCAACCCGA 512
Db 220 GGAGATAAGTCTGTTGGGATGCCAAGCGAAGAAATCTTGGGAAGAGGCCAACCCGA 161

Qy 513 GGACACGAGCTGCTAAATCTCTTTTCTTGATCTCTCTCGGCTGTACCAGACGTAC 572
Db 160 GGACACGAGCTGCTAAATCTCTTTTCTTGATCTCTCTCGGCTGTACCAGACGTAC 101

Qy 573 CGGAGACCAAGCCGACAGATTTCATCCCTTGGATCAGTGAAGCTTGATGTCAGTCCAATAA 632
Db 100 CGGAGACCAAGCCGACAGATTTCATCCCTTGGATCAGTGAAGCTTGATGTCAGTCCAATAA 41

Qy 633 AGCTCCAGCGCAACAAAGACCCCAAGTGAGGTGACACTT 672
Db 40 AGCTCCAGCGCAACAAAGACCCCAAGTGAGGTGACACTT 1

RESULT 10
US-09-513-999C-14399
; Sequence 14399, Application US/0951399C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-12,487
; PRIOR FILING DATE: 1999-02-26
```

```
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14399
; LENGTH: 297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 223
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 237
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 267
; OTHER INFORMATION: k=g or t
US-09-513-999C-14399
```

```
Query Match 7.8%; Score 238; DB 4; Length 297;
Best Local Similarity 98.0%; Pred. No. 4.5e-53;
Matches 249; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 2715 CATTGTGTTAAATTCATCTCAAGGAGGAGAACCCGGCTGTGTGGGTGGTGGCAA 2774
Db 1 CATTGTGTTAAATTCATCTCAAGGAGGAGAACCCGGCTGTGTGGGTGGTGGCAA 60

Qy 2775 TTTCTAGAACGGAATCTGTGGGTATAGAAAAGGAATGAATAAGCGTTGTTTTTCAA 2834
Db 61 TTTCTAGAACGGAATCTGTGGGTATAGAAAAGGAATGAATAAGCGTTGTTTTTCAA 120

Qy 2835 TAGGTCCTTGTAAATTTATTG-ATGAGAGGAGAAAAGATTGACTGGGAGGCTTAAATG 2893
Db 121 TAGGTCCTTGTAAATTTATTG-CATGAGAGGAGAAAAGATTGACTGGGAGGCTTAAATG 180

Qy 2894 ATTTGGAAAAACAATTGCTTTTGGAGCTCAGTGAACCGCAAGATTACACTTAAAAA 2953
Db 181 ATTTGGAAAAACAATTGCTTTTGGAGCTCAGTGAACCGCAAGATTACACTTAAAAA 240

Qy 2954 AAAAAAAAAAAAAA 2967
Db 241 AAAAAAAAAATAA 254
```

RESULT 11

```
US-09-949-016-16018/c
; Sequence 16018, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16018
; LENGTH: 72128
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16018
```

```
Query Match 7.7%; Score 236.8; DB 4; Length 72128;
Best Local Similarity 78.6%; Pred. No. 1.4e-51;
Matches 283; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

| | | | |
|----|-------|---|-------|
| Qy | -2251 | ATGATCACCAAGCAGGAAAGGGAGGCAATAGAGGAATGTGTTCGGGTTAAGTGATGAAAA | 2310 |
| | | | |
| Db | 14607 | ATTGTATCCAAAGTCTTTTAAGCATATAAATCTAAGGAAATGGTATGTAGTAATTTAAAAAGT | 14548 |
| | | | |
| Qy | 2311 | TGGCAGTGGTGGCGCGCGTGGGCTCTCGCCTGTAACTCAGCACATTTGGGAGGCCGA | 2370 |
| | | | |
| Db | 14547 | TGTCTTTGGTGGCTGGAGCGTGGTGGCTCTTGCTGCTATCCAGACATTTGGGAGGCTGA | 14488 |
| | | | |
| Qy | 2371 | GGCAGTGGATCACTTGAGGTCAGGATTTCAAGACTAGCCTGGCCAAACATCATGAACCC | 2430 |
| | | | |
| Db | 14487 | GGCGGCGAGATCACTTGAGCTCAGGATTTGCAGACAGCGCTGGCCAAACATGGTGAAACCT | 14428 |
| | | | |
| Qy | 2431 | CGTCTCTACTTAAAAATACAAAAATTTAGCCAGGCATGGTGGCACACACTGTAGTCCGACG | 2490 |
| | | | |
| Db | 14427 | TGTCCTCTACTTAAAAATACAAAAATTTAGCTGGGCATGGTGACACACACTGTAGTCCCGACG | 14368 |
| | | | |
| Qy | 2491 | TACTCGGGAGCCCAACGCGACAGAAACCGCTTGTACCAGGAGGTGGAGGTTGCAGTGAGC | 2550 |
| | | | |
| Db | 14367 | TACTCGGGAAGCTGAAGCAGGAGAAATCACTTGAAACCCAGAGGTGGAGTTGCAGTGAGC | 14308 |
| | | | |
| Qy | 2551 | CGAAGTTGCACCATTTGCACCTCCACCTCGGGGACAGAGCAAGATTCTTATCAAAAAA | 2610 |
| | | | |
| Db | 14307 | TGAGATCACACCACTGCACCTCAGCCCTGGGTGACAGAGTAAGACTCTGTCACAAAAAA | 14248 |
| | | | |

RESULT 12

```

US-09-949-016-15924/c
; Sequence 15924, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15924
; LENGTH: 42053
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(42053)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15924

```

| | Query Match | 7.6%; | Score 234; | DB 4; | Length 42053; |
|----|-----------------------|---|--------------------|----------------|-------------------|
| | Best Local Similarity | 80.8%; | Pred. No. 5.9e-51; | | |
| | Matches 273; | Conservative | 0; | Mismatches 65; | Indels 0; Gaps 0; |
| QY | 2319 | GTGGCCGGCGTGGTGCTCTCGCCCTGTAATCTCAGCACTTTGGGAGGCCGAGGCAGGTG | 2378 | | |
| DB | 11484 | GGGGTCAGGCGTGGTGGCTCACGCCCTGTAATCCCAAGCACTTTGGGAGGCTGAGCGGGTG | 11425 | | |
| QY | 2379 | GATCACTCAGGTCAGGAGTTCAAGACTAGCCTGGCCCAACATCATGAACCCCGTCTCTA | 2438 | | |
| DB | 11424 | GATCATCTGGGTCAGGAGTTCAAGNCAGCCTGGCCCAACATGGTGAACCCCGTCTCTA | 11365 | | |
| QY | 2439 | CTAAAAATCAAAAAATTACGCAAGGCATGTGGCAACACTGTGTAGTCTCCAGCTACTCGGG | 2498 | | |
| DB | 11364 | CTAAAAATCAAAAAATTACCAAGTCATGTGTGGCAGTGCGCTGTAAATCCAGCTACTCGAG | 11305 | | |
| QY | 2499 | AGCCCCAACCAACAGAACCGCTTGTATCCAGGAGGTGGAGGTTCAGTGTAGCCCAAGTTG | 2558 | | |

| | | | |
|----|-------|---|-------|
| Db | 11304 | AGGCCGAGCAGGAGNAATGGCTTGAACCCAGCAGGTGGAGCTTTTCAGTGCACCGGATG | 11245 |
| Qy | 2559 | CACCAATTGCATCTCCACCTCTGGCGGACAGACGCAAGATTCATCAAAAAAAAAAGGCAGTGG | 2618 |
| Db | 11244 | TGCCACCGCACCTCCAGCTTGGGTGACAGACGAAGACTCCGTCTCAAAAAAAAAAAAAA | 11185 |
| Qy | 2619 | CAACTAAGTTATAGAACAGAAATGCTGTAGAGGAAT | 2656 |
| Db | 11184 | AAAGAACTACAGAACCAAGAAAGACTTTTATAGAAGT | 11147 |

RESULT 13

```

US-09-949-016-17404
; Sequence 17404, Application US/09949016
; Patent NO. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17404
; LENGTH: 137753
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17404

```

| Query Match | 7.6%; | Score 232.2; | DB 4; | Length 137753; |
|-----------------------|-----------------|---|-----------|----------------|
| Best Local Similarity | 80.1%; | Pred. No. 3.2e-50; | | |
| Matches 273; | Conservative 0; | Mismatches 68; | Indels 0; | Gaps 0; |
| Qy | 2305 | TGAAATGCGAGTGGCGCGCGTGTGCTCTCGCTCTAATCTCAGCACATTTGGGA | 2364 | |
| Db | 95156 | TTAAATGCGACCAAAGTCCGGCGCGTGGCTACGCTGTATCCAGCACATTTGGGA | 95215 | |
| Qy | 2365 | GGCCGAGGCGAGGTGAGTCACTTGAGGTGAGGAGTTCAAGACTAGCTCTGGCCCAACATCATG | 2424 | |
| Db | 95216 | GGCCGAGGCGAGGAGTCACTTGAGGTGAGGAGTTTGAGACCGCTGGCCCAACATGNTA | 95275 | |
| Qy | 2425 | AAACCCCGTCTCTACTAAATAATACAAAATTTAGCCAGGCATGGTGGCACACACCTGTAGT | 2484 | |
| Db | 95276 | AAACCCCATCTCTACTAAATAATACAAAATTTATCCAGGCGTGGTGGCTTACCTGTAGT | 95335 | |
| Qy | 2485 | CCGAGTACTTCGGGAGCCCAACGACGACGAGACCGCTGTCTACCCAGGAGGTGGAGTTGCA | 2544 | |
| Db | 95336 | CCGAGTACTTCAGGAGGATGAGGCGAGGAGATGGCTTGAACTCGGAGCGGAGGTTGCA | 95395 | |
| Qy | 2545 | GTGAGCCGAAGTTGCAACCATTTGCACTCCACCCCTGGCCGACAGACAAGATTCTATCAAAA | 2604 | |
| Db | 95396 | GTGAGCCGAGATCAATCACTGCACTCCAGCCTGGGTGACAGAGGAGACTCCATCTCAA | 95455 | |
| Qy | 2605 | AAAAAGGCGAGTGGCAAGTAAAGTTTATGAAGAGAAATGCTG | 2645 | |
| Db | 95456 | GAATAAAAAATGGCAACCAAGTGAATTTCAAGAACTAGGCAG | 95496 | |

RESIT.T 14

```

RES001 14
US-09-349-016-201956
; Sequence 201956, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 16:40:17 ; Search time 1788.95 Seconds

(without alignments)
10749.639 Million cell updates/sec

Title: US-08-731-499-12

Perfect score: 3066

Sequence: 1 GGAACAGCTATGACCATGA.....GGTACCAATGCGCCATA 3066

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 3066 | 100.0 | 3066 | 8 | US-08-731-499-12 |
| 2 | 2828.8 | 92.3 | 3418 | 20 | US-10-723-860-6167 |
| 3 | 2813 | 91.7 | 2813 | 17 | US-10-172-118-937 |
| 4 | 2813 | 91.7 | 2813 | 18 | US-10-240-425-385 |
| 5 | 2813 | 91.7 | 2813 | 18 | US-10-342-887-937 |
| 6 | 2813 | 91.7 | 2813 | 20 | US-10-723-860-1812 |
| 7 | 1929.4 | 62.9 | 2020 | 21 | US-10-491-213-85 |
| | | | | | Sequence 12, Appl |
| | | | | | Sequence 6167, Ap |
| | | | | | Sequence 937, App |
| | | | | | Sequence 385, App |
| | | | | | Sequence 937, App |
| | | | | | Sequence 1812, Ap |
| | | | | | Sequence 85, Appl |

| | | | | | | |
|------|--------|------|--------|----|----------------------|--------------------|
| 8 | 1567 | 51.1 | 2105 | 9 | US-09-939-825-15 | Sequence 15, Appl |
| 9 | 1000.6 | 32.6 | 127567 | 22 | US-10-737-082-47 | Sequence 47, Appl |
| 10 | 1000.6 | 32.6 | 127567 | 22 | US-10-765-790-47 | Sequence 47, Appl |
| c 11 | 582.6 | 19.0 | 644 | 16 | US-10-029-386-22674 | Sequence 22674, A |
| c 12 | 398.6 | 13.0 | 478 | 9 | US-09-783-590-10392 | Sequence 10392, A |
| c 13 | 332.8 | 10.9 | 576 | 16 | US-10-029-386-3914 | Sequence 3914, Ap |
| c 14 | 328 | 10.7 | 372 | 21 | US-10-800-322-327 | Sequence 327, App |
| c 15 | 297 | 9.7 | 506 | 16 | US-10-029-386-8953 | Sequence 8953, Ap |
| c 16 | 265 | 8.6 | 265 | 9 | US-09-864-761-28309 | Sequence 28309, A |
| 17 | 260.4 | 8.5 | 495 | 9 | US-09-864-761-11735 | Sequence 11735, A |
| c 18 | 247 | 8.1 | 349 | 13 | US-10-040-739-889 | Sequence 889, App |
| c 19 | 242.6 | 7.9 | 6096 | 14 | US-10-012-6009-132 | Sequence 132, App |
| c 20 | 236 | 7.7 | 181684 | 13 | US-10-087-192-790 | Sequence 790, App |
| c 21 | 233.8 | 7.6 | 58355 | 21 | US-10-484-577-665 | Sequence 665, App |
| c 22 | 233.8 | 7.6 | 58716 | 21 | US-10-741-600-17754 | Sequence 17754, A |
| c 23 | 233.8 | 7.6 | 186591 | 21 | US-10-484-577-682 | Sequence 682, App |
| c 24 | 233.8 | 7.6 | 208648 | 21 | US-10-484-577-663 | Sequence 663, App |
| c 25 | 233.8 | 7.6 | 208648 | 21 | US-10-484-577-668 | Sequence 668, App |
| c 26 | 232.4 | 7.6 | 227246 | 19 | US-10-322-281-314 | Sequence 314, App |
| c 27 | 231.8 | 7.6 | 627 | 13 | US-10-027-632-250385 | Sequence 250385, A |
| c 28 | 231.8 | 7.6 | 627 | 17 | US-10-027-632-250385 | Sequence 250385, A |
| c 29 | 231.6 | 7.6 | 24699 | 9 | US-09-764-877-2419 | Sequence 2419, Ap |
| c 30 | 231.6 | 7.6 | 24699 | 17 | US-10-242-515-2419 | Sequence 2419, Ap |
| c 31 | 231.6 | 7.6 | 414295 | 20 | US-10-719-993-6876 | Sequence 6876, Ap |
| c 32 | 231.2 | 7.5 | 794 | 13 | US-10-027-632-171663 | Sequence 171663, A |
| c 33 | 231.2 | 7.5 | 794 | 17 | US-10-027-632-171663 | Sequence 171663, A |
| c 34 | 231 | 7.5 | 22047 | 20 | US-10-417-375-12 | Sequence 12, Appl |
| c 35 | 231 | 7.5 | 22056 | 19 | US-10-322-281-106 | Sequence 106, App |
| c 36 | 230.8 | 7.5 | 28693 | 21 | US-10-741-600-17761 | Sequence 17761, A |
| c 37 | 230.8 | 7.5 | 54732 | 19 | US-10-322-281-414 | Sequence 414, App |
| c 38 | 230.6 | 7.5 | 383 | 9 | US-09-960-352-10031 | Sequence 10031, A |
| c 39 | 230.4 | 7.5 | 227931 | 17 | US-10-085-117-274 | Sequence 274, App |
| c 40 | 229.8 | 7.5 | 253861 | 19 | US-10-741-601-5611 | Sequence 5611, Ap |
| c 41 | 229.8 | 7.5 | 261817 | 13 | US-10-087-192-2002 | Sequence 2002, Ap |
| c 42 | 229.4 | 7.5 | 91760 | 13 | US-10-087-192-844 | Sequence 844, App |
| c 43 | 229.4 | 7.5 | 379652 | 21 | US-10-481-613-71 | Sequence 71, Appl |
| c 44 | 229.2 | 7.5 | 285020 | 13 | US-10-087-192-1666 | Sequence 1666, Ap |
| c 45 | 229 | 7.5 | 285 | 21 | US-10-956-157-121436 | Sequence 121436, A |

ALIGNMENTS

RESULT 1

US-08-731-499-12
; Sequence 12, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROWMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3066 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..3066
; OTHER INFORMATION: /note="lbi"
US-08-731-499-12

Query Match      100.0%; Score 3066; DB 8; Length 3066;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3066; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAACAGCTATGACCATGATTACGCCAAGCTCGAATTAACCTCTCACTAAGGGAACA 60
DB 1 GGAACAGCTATGACCATGATTACGCCAAGCTCGAATTAACCTCTCACTAAGGGAACA 60
QY 61 AAGCTGAGCTTCCACCCGGTGGCGCGCTCTAGAACTAGTGTGATCCCCGGGCTGCAG 120
DB 61 AAGCTGAGCTTCCACCCGGTGGCGCGCTCTAGAACTAGTGTGATCCCCGGGCTGCAG 120
QY 121 GAATTCGGCAGAGGCTTCCACCGACAGCCAGGCACTGGGAGCAGCAGCTGGAGCCAG 180
DB 121 GAATTCGGCAGAGGCTTCCACCGACAGCCAGGCACTGGGAGCAGCAGCTGGAGCCAG 180
QY 181 GACCCTGTGAGGAGAGCTCCGGGTGACAGGAGGAGCTGAAATATCTCCACAGGGGC 240
DB 181 GACCCTGTGAGGAGAGCTCCGGGTGACAGGAGGAGCTGAAATATCTCCACAGGGGC 240
QY 241 TCAGCAGAGCAATGGGTAAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAAAT 300
DB 241 TCAGCAGAGCAATGGGTAAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAAAT 300
QY 301 GAACCAAGCAGAGACTTACCAGGACAACCGCTCTGCTCTGAAACCGGGTTCCAGTGGTG 360
DB 301 GAACCAAGCAGAGACTTACCAGGACAACCGCTCTGCTCTGAAACCGGGTTCCAGTGGTG 360
QY 361 GTGTGACCCACACAGTTACGACCTTAGAGGAGTTCGACTTGGGAATAAGTGTCAAGACG 420
DB 361 GTGTGACCCACACAGTTACGACCTTAGAGGAGTTCGACTTGGGAATAAGTGTCAAGACG 420
QY 421 GATAATGTGGCCACTTCTTCCCGGAGACCAACGAGAGATAAGTGTGTGGGGATGCCAAC 480
DB 421 GATAATGTGGCCACTTCTTCCCGGAGACCAACGAGAGATAAGTGTGTGGGGATGCCAAC 480
QY 481 GGAAGAATCTTGGGAAGAGGCCAAACCCGAGGCAACAGCTGCTAAATCTCGTTTTC 540
DB 481 GGAAGAATCTTGGGAAGAGGCCAAACCCGAGGCAACAGCTGCTAAATCTCGTTTTC 540
QY 541 TTGATGCTCTCTGGCTGTACAGGAGCTACCGGAGACCAAGCCGAGATTCATCCCTT 600
DB 541 TTGATGCTCTCTGGCTGTACAGGAGCTACCGGAGACCAAGCCGAGATTCATCCCTT 600
QY 601 GGATCAGTGAAGCTTGTGATGTGAGCTTCCAAATAAGCTCCAGCAACAAAGACCCCAAGTGAG 660
DB 601 GGATCAGTGAAGCTTGTGATGTGAGCTTCCAAATAAGCTCCAGCAACAAAGACCCCAAGTGAG 660
QY 661 AGCTGGACACTTCCGGTGGAGCTGGACCGGGGAGGACACAGATAAAACCCCAAGGAC 720
DB 661 AGCTGGACACTTCCGGTGGAGCTGGACCGGGGAGGACACAGATAAAACCCCAAGGAC 720

661 AGCTGGACACTTCCGGTGGAGCTGGACCGGGGAGGACACAGATAAAACCCCAAGGAC 720
721 GCCCGCGCCCAAGACCAAGGTCCTCTCTGCCCGCAGGATCCCAAGCTTCTCCACCTGAG 780
721 GCCCGCGCCCAAGACCAAGGTCCTCTCTGCCCGCAGGATCCCAAGCTTCTCCACCTGAG 780
781 ACAGGGGAGCAGAGGAGAGCTCCCTCCAAAGCCCAAGGACTCCAGCTTTTGTGACAA 840
781 ACAGGGGAGCAGAGGAGAGCTCCCTCCAAAGCCCAAGGACTCCAGCTTTTGTGACAA 840
841 TTCTTCAAGCTGACCAAGGACAGGAAAGTCCAGGTGACAGCCCAAGGAGCCCAAG 900
841 TTCTTCAAGCTGACCAAGGACAGGAAAGTCCAGGTGACAGCCCAAGGAGCCCAAG 900
901 AGGCGAGAGCATCAAGACAAAGTGGATGAGTTCTTGGCTTATCAGGGCAGTCCGATGAT 960
901 AGGCGAGAGCATCAAGACAAAGTGGATGAGTTCTTGGCTTATCAGGGCAGTCCGATGAT 960
961 GTCCCTGCAGGGAAGGACATAGTTGACGGCAAGGAAAGGACCAAGAACTTTGGAAC 1020
961 GTCCCTGCAGGGAAGGACATAGTTGACGGCAAGGAAAGGACCAAGAACTTTGGAAC 1020
1021 GCGGATGCTCTGCTCCCTGGGACCCAGAGGACTGGAGACTGCAAGAGGACGATTCAC 1080
1021 GCGGATGCTCTGCTCCCTGGGACCCAGAGGACTGGAGACTGCAAGAGGACGATTCAC 1080
1081 GCAGCAGCTATACAGAGATAATAATTCATCATGAGTTCTTTAAACTCTGGTTTCA 1140
1081 GCAGCAGCTATACAGAGATAATAATTCATCATGAGTTCTTTAAACTCTGGTTTCA 1140
1141 CCTAACAAAGCTGAAACCAAAAGGACCCAGAGAGACCGGGTCTGAAAAAGTCAACC 1200
1141 CCTAACAAAGCTGAAACCAAAAGGACCCAGAGAGACCGGGTCTGAAAAAGTCAACC 1200
1201 ACTTACAGTACCTTAAAGTACAGAAAGCCAACTTTACATCCAGGAGACCCAAAGGGCT 1260
1201 ACTTACAGTACCTTAAAGTACAGAAAGCCAACTTTACATCCAGGAGACCCAAAGGGCT 1260
1261 GGCAGAAATTCGAAAGGATGCAACCCATCGGGGACACACAGTCCGTCGACAAACCCCTGAA 1320
1261 GGCAGAAATTCGAAAGGATGCAACCCATCGGGGACACACAGTCCGTCGACAAACCCCTGAA 1320
1321 GCTGCGAAGGAGGACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
1321 GCTGCGAAGGAGGACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
1381 TTTTGGAAAAAGTCAGTTAAAGAGGACTCAGTCCCAAGTGGGAGGAGAAATGTTG 1440
1381 TTTTGGAAAAAGTCAGTTAAAGAGGACTCAGTCCCAAGTGGGAGGAGAAATGTTG 1440
1441 TGTGAGTCACACAGTAGAGATTATAAAGTCCAAAGGAGTAGAATCAGCTTTACAAACAGT 1500
1441 TGTGAGTCACACAGTAGAGATTATAAAGTCCAAAGGAGTAGAATCAGCTTTACAAACAGT 1500
1501 GACCTCAACGAGGAGATGCTGACCTGAAACCCCAAGAGGAAACTCAAAAGAGAGAA 1560
1501 GACCTCAACGAGGAGATGCTGACCTGAAACCCCAAGAGGAAACTCAAAAGAGAGAA 1560
1561 AGCAAAACCAAGAACCTCTCTGATGGGCTTTCTCAGCAAAATGTCAGTGAAGGGATGGA 1620
1561 AGCAAAACCAAGAACCTCTCTGATGGGCTTTCTCAGCAAAATGTCAGTGAAGGGATGGA 1620
1621 GGATCAGCCCACTCAGAAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCC 1680
1621 GGATCAGCCCACTCAGAAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCC 1680
1681 ACAGAAAAAGCTATCAACCGCCAGAGCTTGAAACCAACAGGAGCACACAGAGGGTAAA 1740
1681 ACAGAAAAAGCTATCAACCGCCAGAGCTTGAAACCAACAGGAGCACACAGAGGGTAAA 1740
1741 GAGGGCTCTCGAAGGACAGAGTCAAGCAGCGAGTGAACCAAGGAGGAGCAACAG 1800
1741 GAGGGCTCTCGAAGGACAGAGTCAAGCAGCGAGTGAACCAAGGAGGAGCAACAG 1800
```



```
Qy 1801 CAGAGCCAAAGAACAGCCAGTGCACAGACAGGCCAGCGTGGACAGAACTCACTG 1860
Db 1801 CAGAGCCAAAGAACAGCCAGTGCACAGACAGGCCAGCGTGGACAGAACTCACTG 1860
Qy 1861 CAGAAATGGGCAAGCTCCAAAGAGAGCTCAGAAAGCGGACAGTCCCTTGGGGGCTTC 1920
Db 1861 CAGAAATGGGCAAGCTCCAAAGAGAGCTCAGAAAGCGGACAGTCCCTTGGGGGCTTC 1920
Qy 1921 TTTAAAGCCCTGGGACCAAAGCGGATGTTGATGCTCAAGTGCAAAACAGACCCAGTATCC 1980
Db 1921 TTTAAAGCCCTGGGACCAAAGCGGATGTTGATGCTCAAGTGCAAAACAGACCCAGTATCC 1980
Qy 1981 ATCGACAGCTTGGCAACCAAGCTTAACAATCAGACAGCTTCCACACAGGTTCTCTG 2040
Db 1981 ATCGACAGCTTGGCAACCAAGCTTAACAATCAGACAGCTTCCACACAGGTTCTCTG 2040
Qy 2041 CCACCAAGATGTTCTCTTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Db 2041 CCACCAAGATGTTCTCTTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Qy 2101 TTCTGATGGCCAGCAAAATGAATTCGCTAGAAATTAAGCCGAGCTGTTGTATATGA 2160
Db 2101 TTCTGATGGCCAGCAAAATGAATTCGCTAGAAATTAAGCCGAGCTGTTGTATATGA 2160
Qy 2161 GGTGATATTTACGCTCTCTGTCAGTCTTTCTTCTGCAATTAACAGTAAAGTGGTTTA 2220
Db 2161 GGTGATATTTACGCTCTCTGTCAGTCTTTCTTCTGCAATTAACAGTAAAGTGGTTTA 2220
Qy 2221 GCAGTCACTAGTGGTTCAGAGAGTTCGATGATCAACAGCAGGAAGGAGGAATA 2280
Db 2221 GCAGTCACTAGTGGTTCAGAGAGTTCGATGATCAACAGCAGGAAGGAGGAATA 2280
Qy 2281 GAGAAATGTGTTGGGTTAAGTGTAAATGAGAGTGTGGCGGGCGTGGTGGCTCTC 2340
Db 2281 GAGAAATGTGTTGGGTTAAGTGTAAATGAGAGTGTGGCGGGCGTGGTGGCTCTC 2340
Qy 2341 GCCTGTAATCTCAGCACTTTGGGAGCGCGAGGAGGTGGATCACTGAGGTCAAGAGTTC 2400
Db 2341 GCCTGTAATCTCAGCACTTTGGGAGCGCGAGGAGGTGGATCACTGAGGTCAAGAGTTC 2400
Qy 2401 AAGACTAGCTGGCCAAATCATCAATCAAAACCCGCTCTCTACTAAATAACAAATAAGCCA 2460
Db 2401 AAGACTAGCTGGCCAAATCATCAATCAAAACCCGCTCTCTACTAAATAACAAATAAGCCA 2460
Qy 2461 GGCAATGTTGGCACAACCTGTAGTCCCAGCTACTCGGAGGCCCAACGACGAGAACGCT 2520
Db 2461 GGCAATGTTGGCACAACCTGTAGTCCCAGCTACTCGGAGGCCCAACGACGAGAACGCT 2520
Qy 2521 TGTACCCAGGAGGTGGAGGTTGCAGTGAGCGGAAGTTGCACATTCGCACTCCACCTGGG 2580
Db 2521 TGTACCCAGGAGGTGGAGGTTGCAGTGAGCGGAAGTTGCACATTCGCACTCCACCTGGG 2580
Qy 2581 CGACAGACAGATTTCTATCAAAAATAAGGAGTGCAGTAAGTTATAGAGAGAA 2640
Db 2581 CGACAGACAGATTTCTATCAAAAATAAGGAGTGCAGTAAGTTATAGAGAGAA 2640
Qy 2641 TGCTGCTAGAGGAATTAAGCGTTGTAGTAAACCGTGTCTCATCTCTAAGCTTGAAGAA 2700
Db 2641 TGCTGCTAGAGGAATTAAGCGTTGTAGTAAACCGTGTCTCATCTCTAAGCTTGAAGAA 2700
Qy 2701 GGGAGACGAAATCCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGT 2760
Db 2701 GGGAGACGAAATCCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGT 2760
Qy 2761 TGGTGTGTCGCAATTTCTTAGACGGAATGTGGGTATAGAAAAGGATGAATAAG 2820
Db 2761 TGGTGTGTCGCAATTTCTTAGACGGAATGTGGGTATAGAAAAGGATGAATAAG 2820
Qy 2821 CGTTGTTTTTCAAAATAGGCTCTTGTAAAGTTATTTGATGAGAGGAGAAAGATTGACTGGG 2880
Db 2821 CGTTGTTTTTCAAAATAGGCTCTTGTAAAGTTATTTGATGAGAGGAGAAAGATTGACTGGG 2880
```

```
Qy 2881 AGGCTTAAATGATTTGGGAAACAAATTCGTTTTTGGGCTCACTGACAAACGGCAAGAT 2940
Db 2881 AGGCTTAAATGATTTGGGAAACAAATTCGTTTTTGGGCTCACTGACAAACGGCAAGAT 2940
Qy 2941 TACAACCTTAAATAAAAAAAAAAAAAAACTCGAGACTAGTTCTCTCTCTCTCTCTCTCGTCCGA 3000
Db 2941 TACAACCTTAAATAAAAAAAAAAAAAAACTCGAGACTAGTTCTCTCTCTCTCTCTCTCGTCCGA 3000
Qy 3001 ATTGATATCAAGCTTATCGATACCGTCCGACTCGAGGGGGGGCCCGGTACCCAAATTCGC 3060
Db 3001 ATTGATATCAAGCTTATCGATACCGTCCGACTCGAGGGGGGGCCCGGTACCCAAATTCGC 3060
Qy 3061 CCTATA 3066
Db 3061 CCTATA 3066

RESULT 2
US-10-723-860-6167
; Sequence 6167, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6167
; LENGTH: 3418
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1815)..(1815)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3303)..(3416)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-6167
```

Query Match 92.3%; Score 2828.8; DB 20; Length 3418;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2833; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
Qy 127 GGACACGAGGCTCCACACGACGAGCCAGGCACTGGGCGAGCACGACCTGGAGACCCAGACCCCT 186
Db 40 GGACACGAGGCTCCACACGAGGCTCCAGGCACTGGGCGAGCACGACCTGGAGACCCAGACCCCT 99
Qy 187 GTGCAGGAGCAGCTCCCGGTGACACGAGGGGACTGAAGATACTCCCAAGGGGCTCAGCA 246
Db 100 GTGCAGGAGCAGCTCCCGGTGACACGAGGGGACTGAAGATACTCCCAAGGGGCTCAGCA 159
Qy 247 GGAGCAATGGGTAAACCAATGAGTGTCCCAAGAGTGTGAAGACCAAGAGATGAACCA 306
Db 160 GGAGCAATGGGTAAACCAATGAGTGTCCCAAGAGTGTGAAGACCAAGAGATGAACCA 219
Qy 307 GAAGCAGAGACTTACACGAGCAACGCGTCTGCTGAACCGGGTCCAGTGGTGGTGTGCG 366
Db 220 GAAGCAGAGACTTACACGAGCAACGCGTCTGCTGAACCGGGTCCAGTGGTGGTGTGCG 279
Qy 367 ACCCACAAGTTCAGCACTTAGAGAAAGTGCAGTTGGGAATAAGTGTCAAGACCGGATAAT 426
Db 280 ACCCACAAGTTCAGCACTTAGAGAAAGTGCAGTTGGGAATAAGTGTCAAGACCGGATAAT 339
Qy 427 GTGGCCACTTCTTCTCCCGGAGACAAACGAGAGATAAGTGTGTTCGGGATGCCAACGGAAAG 486
```

Db 340 GTGGCCACTTCTTCCCGGAGACAAACCGAGATAAGTCTGTGGAGTGCACCGGAAAG 399
Qy 487 AATCTTTGGGAAGAGGCCAAACCCGAGGACCCAGCTCTTAATCTCGTTTCTTTTGATG 546
Db 400 AATCTTTGGGAAGAGGCCAAACCCGAGGACCCAGCTCTTAATCTCGTTTCTTTTGATG 459
Qy 547 CTCTCTCGGCTGTACACAGAGAGTACCGGAGACCAAGCCGAGATTCATCCCTTGGATCA 606
Db 460 CTCTCTCGGCTGTACACAGAGAGTACCGGAGACCAAGCCGAGATTCATCCCTTGGATCA 519
Qy 607 GTGAAGCTTGATGTACGCTCCAATAAAGCTCCAGCGAACCAAGACCCCAAGTGAGAGCTGG 666
Db 520 GTGAAGCTTGATGTACGCTCCAATAAAGCTCCAGCGAACCAAGACCCCAAGTGAGAGCTGG 579
Qy 667 ACATTTCCGCTGGCAGCTGGACCGGGGACGACACAGATTAATACCCCGAGGGCAGCCCGG 726
Db 580 ACATTTCCGCTGGCAGCTGGACCGGGGACGACACAGATTAATACCCCGAGGGCAGCCCGG 639
Qy 727 GCCCAAGACAAAGTCTCTCTGCGCCGAGGATCCCAAGCTTCTCCACCTGACAGAGGG 786
Db 640 GCCCAAGACAAAGTCTCTCTGCGCCGAGGATCCCAAGCTTCTCCACCTGACAGAGGG 699
Qy 787 GGAGCAGGAGAGAGTCTCCTCCAAAGCTCCAGGATCCAGCTTTTGTGACAAATCTTC 846
Db 700 GGAGCAGGAGAGAGTCTCCTCCAAAGCTCCAGGATCCAGCTTTTGTGACAAATCTTC 759
Qy 847 AAGCTGACAAAGGACAGGAAAGGTGCCAGGTGACAGCCCAAGGAGCCNAGAGGGCA 906
Db 760 AAGCTGACAAAGGACAGGAAAGGTGCCAGGTGACAGCCCAAGGAGCCNAGAGGGCA 819
Qy 907 GAGCATCAAGACAAAGTGGATGAGTTCTTGCTTATCAGGGCAGTCCGATGATGCTCCT 966
Db 820 GAGCATCAAGACAAAGTGGATGAGTTCTTGCTTATCAGGGCAGTCCGATGATGCTCCT 879
Qy 967 GCAGGGAAGGACATAGTTGACCGCAAGGAAAGGAGCAAGAACTTGGAACTTGCAGAT 939
Db 1027 TGCTCTCTCCTGGGACCCAGAGGACTGGAGACTCCNAGGACGATTCACCGGACGCA 1086
Qy 940 TGCTCTCTCCTGGGACCCAGAGGACTGGAGACTCCNAGGACGATTCACCGGACGCA 999
Qy 1087 GCTATAGCAGAGATAATATTCATCATGAGTTCTTTTAAACTCTGGTTTACACTTAAC 1146
Db 1000 GCTATAGCAGAGATAATATTCATCATGAGTTCTTTTAAACTCTGGTTTACACTTAAC 1059
Qy 1147 AAAGCTGAAACAAAGGACCCAGAGACACCGGGTGTGAAAAGTCAACCCACACTTCA 1206
Db 1060 AAAGCTGAAACAAAGGACCCAGAGACACCGGGTGTGAAAAGTCAACCCACACTTCA 1119
Qy 1207 GCTGACCTTAAGTCAGACNAGCCAACTTTTACATCCAGGAGACCCAGGGCTGGCAAG 1266
Db 1120 GCTGACCTTAAGTCAGACNAGCCAACTTTTACATCCAGGAGACCCAGGGCTGGCAAG 1179
Qy 1267 AATTCCAAAGATGCAACCCATCCGGGACACACAGTCCGTCACAAACCCCTGAACTTGG 1326
Db 1180 AATTCCAAAGATGCAACCCATCCGGGACACACAGTCCGTCACAAACCCCTGAACTTGG 1239
Qy 1327 AAGGAAGCCACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGG 1386
Db 1240 AAGGAAGCCACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGG 1299
Qy 1387 AAAAGTCAGTTAAGAGGACTAGTCCCAAGTGGGAGGAGATGTGTTGTGAG 1446
Db 1300 AAAAGTCAGTTAAGAGGACTAGTCCCAAGTGGGAGGAGATGTGTTGTGAG 1359
Qy 1447 TCACAGTAGAGATTATTAAGTCCAAAGGAGTGAATCAGCTTTACAAACAGTGGACCTC 1506
Db 1360 TCACAGTAGAGATTATTAAGTCCAAAGGAGTGAATCAGCTTTACAAACAGTGGACCTC 1419
Qy 1507 AACGAAGAGATGCTGCACTTGAAACCCACAGAGCGGAAACTCAAAAGAGAAAGCAAA 1566

Db 1420 AACGAAGGAGATGCTGCACCTGAACCCACAGAACCGAAACTCAAAAGAGAGAAAGCAAA 1479
Qy 1567 CCAAGAACCTCTCTGATGGGTTTCTCAGACAAATGTTCAGTGAAGGGGATGAGGGATC 1626
Db 1480 CCAAGAACCTCTCTGATGGGTTTCTCAGACAAATGTTCAGTGAAGGGGATGAGGGATC 1539
Qy 1627 ACCCACTCAGAAGAAATAATGGGAAGACTCCAGCTGCCAAAATCAGACTCCACAGAA 1686
Db 1540 ACCCACTCAGAAGAAATAATGGGAAGACTCCAGCTGCCAAAATCAGACTCCACAGAA 1599
Qy 1687 AAGACTATCACACCGCAGAGCTGAACCAACAGGAGCACCACAGAGGGTAAAGGGC 1746
Db 1600 AAGACTATCACACCGCAGAGCTGAACCAACAGGAGCACCACAGAGGGTAAAGGGC 1659
Qy 1747 TCCTCGAAGGACAAAGAGTCCAGCAGCCGAGATGAACAAAGCAGAGACAAACAGCAGAA 1806
Db 1660 TCCTCGAAGGACAAAGAGTCCAGCAGCCGAGATGAACAAAGCAGAGACAAACAGCAGAA 1719
Qy 1807 GCCAAGAACCCAGCTCCAGTGCACAGAGCAGGCCACCGTGGACACGAACTCACTGCAAT 1866
Db 1720 GCCAAGAACCCAGCTCCAGTGCACAGAGCAGGCCACCGTGGACACGAACTCACTGCAAT 1779
Qy 1867 GGGGACAGCTCCAAAAGAGACCTGAGAGCGCAGAGTCCCTTGGGGGCTTTTAAA 1926
Db 1780 GGGGACAGCTCCAAAAGAGACCTGAGAGCGCAGAGTCCCTTGGGGGCTTTTAAA 1839
Qy 1927 GGCCTGGGACCAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGA 1986
Db 1840 GGCCTGGGACCAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGA 1899
Qy 1987 CCAGTTGGCAAAACCCAGTAAACAAATCAGACCGTTCCACCAGGTTTCTCTGCCACCA 2046
Db 1900 CCAGTTGGCAAAACCCAGTAAACAAATCAGACCGTTCCACCAGGTTTCTCTGCCACCA 1959
Qy 2047 AGATGTTCTCTCTTACTCCATCTCTCTCCCAACACGCTCCATGATATATTTCTTGA 2106
Db 1960 AGATGTTCTCTCTTACTCCATCTCTCTCCCAACACGCTCCATGATATATTTCTTGA 2019
Qy 2107 TGCCAGCAAAATGAAATTTCTGCTTAGAAAATAAGCCCGAGCTGTTGATATTTAGAGTGTA 2166
Db 2020 TGCCAGCAAAATGAAATTTCTGCTTAGAAAATAAGCCCGAGCTGTTGATATTTAGAGTGTA 2079
Qy 2167 TTATTTACGTCCTCTGCTCCAGTCTTTTCTGGCAAAATAACAGTAAAGTGTATAGCAGGT 2226
Db 2080 TTATTTACGTCCTCTGCTCCAGTCTTTTCTGGCAAAATAACAGTAAAGTGTATAGCAGGT 2139
Qy 2227 CACCTAGTTGGTTCAGNAGAGTCGATGATCACCAGCAGAGAGGGAGGAAATAGAGAA 2286
Db 2140 CACCTAGTTGGTTCAGNAGAGTCGATGATCACCAGCAGAGAGGGAGGAAATAGAGAA 2199
Qy 2287 TGTGTTCCGGTTAAGTGCATGAAAATGGCAGTGGTGGCCGGCGTGGTCTCTGCCCTGT 2346
Db 2200 TGTGTTCCGGTTAAGTGCATGAAAATGGCAGTGGTGGCCGGCGTGGTCTCTGCCCTGT 2259
Qy 2347 AATCTCAGCACTTTGGGAGGCCGAGGAGTGGATCACCCTGAGGTCAAGGATTCAGAGT 2406
Db 2260 AATCTCAGCACTTTGGGAGGCCGAGGAGTGGATCACCCTGAGGTCAAGGATTCAGAGT 2319
Qy 2407 AGCTGGCCACATCATGAAAACCCCGTCTCTACTTAAATAATACAAAATTTAGCCAGGATG 2466
Db 2320 AGCTGGCCACATCATGAAAACCCCGTCTCTACTTAAATAATACAAAATTTAGCCAGGATG 2379
Qy 2467 GTGSCACACACCTGTAGTCCCGACTCTCGGGAGCCCAACGACAGAAACCGCTTGTACC 2526
Db 2380 GTGSCACACACCTGTAGTCCCGACTCTCGGGAGCCCAACGACAGAAACCGCTTGTACC 2439
Qy 2527 CAGGAGTGGAGTTGAGTGCAGCCGAAAGTTGCAACCATTTGCACTCCACCTTGGGGCAG 2586
Db 2440 CAGGAGTGGAGTTGAGTGCAGCCGAAAGTTGCAACCATTTGCACTCCACCTTGGGGCAG 2499
Qy 2587 AGCAGAGTTCTATCAAAAAAAGGAGCAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGC 2646
Db 2500 AGCAGAGTTCTATCAAAAAAAGGAGCAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGC 2559


```
QY 1456 GAGATTATAAGTCCAAAGGAAGTAGAATCAGCCCTTACAAACAGTGGACCTCAACGAAGGA 1515
DB |||||
QY 1321 GAGATTATAAGTCCAAAGGAAGTAGAATCAGCCCTTACAAACAGTGGACCTCAACGAAGGA 1380
DB |||||
QY 1516 GATGCTGCACCTGAACCCACAGAAAGCGAAACTCAAAAGAGAAAGAAACCAACGAACCC 1575
DB |||||
QY 1381 GATGCTGCACCTGAACCCACAGAAAGCGAAACTCAAAAGAGAAAGAAACCAACGAACCC 1440
DB |||||
QY 1576 TCTCTGATGGCTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGGATCACCCTCA 1635
DB |||||
QY 1441 TCTCTGATGGCTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGGATCACCCTCA 1500
DB |||||
QY 1636 GAAGAAATAAATGGGAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGACTATC 1695
DB |||||
QY 1501 GAAGAAATAAATGGGAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGACTATC 1560
DB |||||
QY 1696 ACACCGCCAGAGCTGAACCAACAGGAGCACACAGAAAGGGTAAAGAGGCTCCTCGAAG 1755
DB |||||
QY 1561 ACACCGCCAGAGCTGAACCAACAGGAGCACACAGAAAGGGTAAAGAGGCTCCTCGAAG 1620
DB |||||
QY 1756 GACAAAGATCAGACAGCGGAGTGAACAAAGCAGAGAGCAACAGCAGGAAGCCAAAGAA 1815
DB |||||
QY 1621 GACAAAGATCAGACAGCGGAGTGAACAAAGCAGAGAGCAACAGCAGGAAGCCAAAGAA 1680
DB |||||
QY 1816 CCAGCCAGTGCCACAGAGCGGCCACCGTGGACACGAACTCCTGCAGAAATGGGGACAAG 1875
DB |||||
QY 1681 CCAGCCAGTGCCACAGAGCGGCCACCGTGGACACGAACTCCTGCAGAAATGGGGACAAG 1740
DB |||||
QY 1876 CTCCAAAGAGACTGAGAGAGCGGCAGAGTCCCTTGGGGCTTCTTTAAAGCGCTGGGA 1935
DB |||||
QY 1741 CTCCAAAGAGACTGAGAGAGCGGCAGAGTCCCTTGGGGCTTCTTTAAAGCGCTGGGA 1800
DB |||||
QY 1936 CCAAGCGGATGTTGGATGCTCAAGTCCAAACAGACCCAGTATCCATCCGACCAAGTGGC 1995
DB |||||
QY 1801 CCAAGCGGATGTTGGATGCTCAAGTCCAAACAGACCCAGTATCCATCCGACCAAGTGGC 1860
DB |||||
QY 1996 AAACCCAAAGTAAACAAATCAGACGGTTCCTCCACAGAGTTCCTCCGACCAAGATGTT 2055
DB |||||
QY 1861 AAACCCAAAGTAAACAAATCAGACGGTTCCTCCACAGAGTTCCTCCGACCAAGATGTT 1920
DB |||||
QY 2056 CTCCTTACTCCATCTCTCCCAACACAGCTCCATGATATATATTTCTGATGCCAGCA 2115
DB |||||
QY 1921 CTCCTTACTCCATCTCTCCCAACACAGCTCCATGATATATATTTCTGATGCCAGCA 1980
DB |||||
QY 2116 AATGAATTCCTGCTAGAAATTAAGCCGAGCTGCTGATATATTTGAGGTGATTTATTTACG 2175
DB |||||
QY 1981 AATGAATTCCTGCTAGAAATTAAGCCGAGCTGCTGATATATTTGAGGTGATTTATTTACG 2040
DB |||||
QY 2176 TCTCTGCTCCAGTCTTTTCTGGCAAATAACAGTAAAGATGGTTTTCAGAGGTTCACCTAGTT 2235
DB |||||
QY 2041 TCTCTGCTCCAGTCTTTTCTGGCAAATAACAGTAAAGATGGTTTTCAGAGGTTCACCTAGTT 2100
DB |||||
QY 2236 GGGTCAGAAAGTCCGATGATCACAAGCAGAAAGGGAGGGAATAGAGGAATGTTTCGG 2295
DB |||||
QY 2101 GGGTCAGAAAGTCCGATGATCACAAGCAGAAAGGGAGGGAATAGAGGAATGTTTCGG 2160
DB |||||
QY 2296 GTTAAGTGAATAAATGGCAGTGGTGGCCGGCGTGGTGGCTCTCGCTGTAATCTCAGC 2355
DB |||||
QY 2161 GTTAAGTGAATAAATGGCAGTGGTGGCCGGCGTGGTGGCTCTCGCTGTAATCTCAGC 2220
DB |||||
QY 2356 ACTTTGGAGGCCGAGGCGAGGTGGATCACCTGAGGTTCAGAGGTTCAGAGGTTCAGAGGTTCAGG 2415
DB |||||
QY 2221 ACTTTGGAGGCCGAGGCGAGGTGGATCACCTGAGGTTCAGAGGTTCAGAGGTTCAGAGGTTCAGG 2280
DB |||||
QY 2416 AACATCATGAACCCCGTCTCTACTAAATAATACAAAATTAAGCAGGCGATGGTGGCACAC 2475
DB |||||
QY 2281 AACATCATGAACCCCGTCTCTACTAAATAATACAAAATTAAGCAGGCGATGGTGGCACAC 2340
DB |||||
QY 2476 ACCTGTAGTCCAGCTACTCGGAGGCCCAACGACGAGAACCGCTTGTACCCAGGAGGTG 2535
DB |||||
QY 2341 ACCTGTAGTCCAGCTACTCGGAGGCCCAACGACGAGAACCGCTTGTACCCAGGAGGTG 2400
DB |||||
```

```
QY 2536 GAGGTTGCACTGAGCCGAAGTTGCAACATTTGCACTCCACCTCGGGCGACAGAGCAAGATT 2595
DB |||||
QY 2401 GAGGTTGCACTGAGCCGAAGTTGCAACATTTGCACTCCACCTCGGGCGACAGAGCAAGATT 2460
DB |||||
QY 2596 CTATCAAAAAAAGGCGAGTGGCAAGTATAGTAGAGAAATCTCTCTAGAGGAA 2655
DB |||||
QY 2461 CTATCAAAAAAAGGCGAGTGGCAAGTATAGTAGAGAAATCTCTCTAGAGGAA 2520
DB |||||
QY 2656 TTAAGCGTTTCTAGTAAACGGCTCTCATCTCTTAAGCTTGAAGAGGGAGACGAAAAATCC 2715
DB |||||
QY 2521 TTAAGCGTTTCTAGTAAACGGCTCTCATCTCTTAAGCTTGAAGAGGGAGACGAAAAATCC 2580
DB |||||
QY 2716 ATTTGTTTAAATTCATCTCAAGAGGGAGAACCCGGGCTGTGTTGGGTGGTTGCCAAT 2775
DB |||||
QY 2581 ATTTGTTTAAATTCATCTCAAGAGGGAGAACCCGGGCTGTGTTGGGTGGTTGCCAAT 2640
DB |||||
QY 2776 TTCTTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAAAT 2835
DB |||||
QY 2641 TTCTTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAAAT 2700
DB |||||
QY 2836 AGGTCTCTTCTAAGTTTATTCATGAGAGGGAAAAAGATTGACTCGGGAGGGCTTAAAAATGAT 2895
DB |||||
QY 2701 AGGTCTCTTCTAAGTTTATTCATGAGAGGGAAAAAGATTGACTCGGGAGGGCTTAAAAATGAT 2760
DB |||||
QY 2896 TTGGGAAAAACAATTTGCTTTTGGGCTCAGTGACAAACGGCAAGATTACAACTT 2948
DB |||||
QY 2761 TTGGGAAAAACAATTTGCTTTTGGGCTCAGTGACAAACGGCAAGATTACAACTT 2813
DB |||||
```

RESULT 4

```
US-10-240-425-385
; Sequence 385, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Bolland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 385
; LENGTH: 2813
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AF041260
US-10-240-425-385
```

```
Query Match 91.7%; Score 2813; DB 18; Length 2813;
Best local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 136 CTCACCGACAGCCAGCCAGTGGSCAGCAGCAGTGGAGACCCAGGACCCCTGTGCAGGAG 195
DB 1 CTCACCGACAGCCAGCCAGTGGSCAGCAGCAGTGGAGACCCAGGACCCCTGTGCAGGAG 60
QY 196 CAGTCCCGGTGACACAGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCAGGAGCAATG 255
DB 61 CAGTCCCGGTGACACAGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCAGGAGCAATG 120
QY 256 GGTAAACCAATGAGTGTTCCTCCAAAGATTGAAGACCAAGAGATGAACCAAGAGCAGAG 315
DB |||||
```

Db 121 GGTAACCAATGAGTGTTCCTCCCAAGAGTGAAGACCAAGAGATGMAACCAAGCAGAG 180
Qy 316 ACTTACCAAGCAACGGCTCTGCTGTAACGGGGTTCAGTGGTGTGTCACCAACACA 375
Db 181 ACTTACCAAGCAACGGCTCTGCTGTAACGGGGTTCAGTGGTGTGTCACCAACACA 240
Qy 376 GTTCAGCACTTAGAGGAAGTGCATCTGGGAATAAGTGTCAAGAGCGGATAATGTGGCCACT 435
Db 241 GTTCAGCACTTAGAGGAAGTGCATCTGGGAATAAGTGTCAAGAGCGGATAATGTGGCCACT 300
Qy 436 TCTTCCCGCAGACAACGGAGATAGTGTCTGTGGGATGCGATGCAACGGGAAGAAATCTTTGGG 495
Db 301 TCTTCCCGCAGACAACGGAGATAAGTGTCTGTGGGATGCGATGCAACGGGAAGAAATCTTTGGG 360
Qy 496 AAAGAGCCAAACCGGAGCAGCTGCTAAATCTCGTTTTTCTTGTGATGCTCTCTCGG 555
Db 361 AAAGAGCCAAACCGGAGCAGCTGCTAAATCTCGTTTTTCTTGTGATGCTCTCTCGG 420
Qy 556 CCTGTACAGGACGTACCGGAGACAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTT 615
Db 421 CCTGTACAGGACGTACCGGAGACAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTT 480
Qy 616 GATCTCAGCTCCAATAAGCTCCAGCGAACAAGACCCCAAGTGAAGCTGGACACTTCCG 675
Db 481 GATCTCAGCTCCAATAAGCTCCAGCGAACAAGACCCCAAGTGAAGCTGGACACTTCCG 540
Qy 676 GTGCAGCTGACCGGGCAGGACACAGATAAAACCCAGGGCAGCCCGGCCCAAGAC 735
Db 541 GTGCAGCTGACCGGGCAGGACACAGATAAAACCCAGGGCAGCCCGGCCCAAGAC 600
Qy 736 AAGTGTCTCTGCGCCGAGGATCCACGCTTCTCCACCTGAGACAGGGGGAGCAGGA 795
Db 601 AAGTGTCTCTGCGCCGAGGATCCACGCTTCTCCACCTGAGACAGGGGGAGCAGGA 660
Qy 796 GGAGAAGCTCCTCCAAAGCCCAAGGACTCCAGCTTTTTTGTGACAAATCTTCAAGCTGGAC 855
Db 661 GGAGAAGCTCCTCCAAAGCCCAAGGACTCCAGCTTTTTTGTGACAAATCTTCAAGCTGGAC 720
Qy 856 AAGGGACAGNAAGTGTCCAGGTGACGCGCAACAGNAGCCAAAGAGCGCAGACATCAA 915
Db 721 AAGGGACAGNAAGTGTCCAGGTGACGCGCAACAGNAGCCAAAGAGCGCAGACATCAA 780
Qy 916 GACAAGTGGATGAGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCTGCAAGGGAAG 975
Db 781 GACAAGTGGATGAGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCTGCAAGGGAAG 840
Qy 976 GACATAGTTGACGGCAAGGAAAGAGGACAGAACTTGGAACTGCGGATGCTCTGTC 1035
Db 841 GACATAGTTGACGGCAAGGAAAGAGGACAGAACTTGGAACTGCGGATGCTCTGTC 900
Qy 1036 CCTGGGACCCAGAAAGTGGAGACTGCAAGAGCGATTCCTCAGGAGCAGCTATAGCA 1095
Db 901 CCTGGGACCCAGAAAGTGGAGACTGCAAGAGCGATTCCTCAGGAGCAGCTATAGCA 960
Qy 1096 GAGATAATTAATTCATCAGTTCTTTTAAACTCTGGTTTCACTTCACTCAAAAGCTGAA 1155
Db 961 GAGATAATTAATTCATCAGTTCTTTTAAACTCTGGTTTCACTTCACTCAAAAGCTGAA 1020
Qy 1156 ACAAAAAGGACCCAGAAAGACACGGGTGCTGAAAAGTCAACCACCACTTCAGCTGACCTT 1215
Db 1021 ACAAAAAGGACCCAGAAAGACACGGGTGCTGAAAAGTCAACCACCACTTCAGCTGACCTT 1080
Qy 1216 AAGTCAGACAAGCCAACTTTATATCCAGAGAGACCAAGGGGTGGCAAGAAATTCAAA 1275
Db 1081 AAGTCAGACAAGCCAACTTTATATCCAGAGAGACCAAGGGGTGGCAAGAAATTCAAA 1140
Qy 1276 GGATGCAACCATCGGGGACACACAGTCCGTGACACCCCTGAACTGCGAAGAGGC 1335
Db 1141 GGATGCAACCATCGGGGACACACAGTCCGTGACACCCCTGAACTGCGAAGAGGC 1200
Qy 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1395
Db 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1260

Qy 1396 GTTAAAGGAGACTCAGTCCCCACAGGTCCGAGGAGAGATGTGGTGTGAGTCACCACTA 1455
Db 1261 GTTAAAGGAGACTCAGTCCCCACAGGTCCGAGGAGAGATGTGGTGTGAGTCACCACTA 1320
Qy 1456 GAGATTATAAGTCCAAAGGAAGTAGAATCAGCTTTCAAAAAGTGGACCTCAACGAGGA 1515
Db 1321 GAGATTATAAGTCCAAAGGAAGTAGAATCAGCTTTCAAAAAGTGGACCTCAACGAGGA 1380
Qy 1516 GATCTGCTCAGCTGAAACCCACAGAGCGAAACTCAAAAAGAGAGAAACAAACCAAGAAC 1575
Db 1381 GATCTGCTCAGCTGAAACCCACAGAGCGAAACTCAAAAAGAGAGAAACAAACCAAGAAC 1440
Qy 1576 TCTCTGATGGCTTTCTCAGACCAATGTCAAGTGAAGGGGATGAGGGATCACCCACTCA 1635
Db 1441 TCTCTGATGGCTTTCTCAGACCAATGTCAAGTGAAGGGGATGAGGGATCACCCACTCA 1500
Qy 1636 GAAGAAATAATGGGAAGACTCCAGCTGCCAAAACATCAGACTCCACAGAAAAGACTATC 1695
Db 1501 GAAGAAATAATGGGAAGACTCCAGCTGCCAAAACATCAGACTCCACAGAAAAGACTATC 1560
Qy 1696 ACACCGCAGAGCTGAAACCAACAGGAGCACCAAGAGGGTAAAGGGCTCTCTCGAAG 1755
Db 1561 ACACCGCAGAGCTGAAACCAACAGGAGCACCAAGAGGGTAAAGGGCTCTCTCGAAG 1620
Qy 1756 GACAAGAAAGTCAAGCGCGAGATGAACAGCAGAAAGACCAACAGCAGGAAGCCAAAGAA 1815
Db 1621 GACAAGAAAGTCAAGCGCGAGATGAACAGCAGAAAGACCAACAGCAGGAAGCCAAAGAA 1680
Qy 1816 CCAGCCAGTGCACAGAGCAGGCCACAGTGACACGAACTCAGCTGCAGAAATGGGACAAAG 1875
Db 1681 CCAGCCAGTGCACAGAGCAGGCCACAGTGACACGAACTCAGCTGCAGAAATGGGACAAAG 1740
Qy 1876 CTCAAAGAGACCTGAGAAAGCGGACAGTCCCTTGGGGCTCTTTAAAGGCTTGGGA 1935
Db 1741 CTCAAAGAGACCTGAGAAAGCGGACAGTCCCTTGGGGCTCTTTAAAGGCTTGGGA 1800
Qy 1936 CCAAGCGGATGTGGATGCTCAAGTGCACACAGCCAGTATCATCTGACACAGTTGGC 1995
Db 1801 CCAAGCGGATGTGGATGCTCAAGTGCACACAGCCAGTATCATCTGACAGTGGC 1860
Qy 1996 AAACCCAAAGTAAACAAATCAGACGGTTCCACCAAGTTCCTGCGCACCAAGATGCTT 2055
Db 1861 AAACCCAAAGTAAACAAATCAGACGGTTCCACCAAGTTCCTGCGCACCAAGATGCTT 1920
Qy 2056 CTCTTACTCCATCTCTCCCAACACGCTCCATATATATATTTCTTCTGATGGCCAGCA 2115
Db 1921 CTCTTACTCCATCTCTCCCAACACGCTCCATATATATATTTCTTCTGATGGCCAGCA 1980
Qy 2116 AATGAAATTCCTGCTGAAATTAAGCCCGAGCTGTTGTATATTTAGAGTGTATTTACG 2175
Db 1981 AATGAAATTCCTGCTGAAATTAAGCCCGAGCTGTTGTATATTTAGAGTGTATTTACG 2040
Qy 2176 TCTCTGTGCTCAGTCTTTTCTGGCAAAATAACAGTAAAGTGGTTTAGCAGGTCACTAGTT 2235
Db 2041 TCTCTGTGCTCAGTCTTTTCTGGCAAAATAACAGTAAAGTGGTTTAGCAGGTCACTAGTT 2100
Qy 2236 GGGTCAGAGAGTGCATGATCAACCAAGCAGAAAGGGAGGAAATAGAGGAATGTGTTGG 2295
Db 2101 GGGTCAGAGAGTGCATGATCAACCAAGCAGAAAGGGAGGAAATAGAGGAATGTGTTGG 2160
Qy 2296 GTTAAAGTGAATAAATGGCAGTGGTGGCCGGCTGTGGTCTCTGCTGTGTAATCTCAGC 2355
Db 2161 GTTAAAGTGAATAAATGGCAGTGGTGGCCGGCTGTGGTCTCTGCTGTGTAATCTCAGC 2220
Qy 2356 ACTTTGGGAGCCGAGGACAGTGGATCACCTGAGTCAAGGATTCAGAGCTAGCTGCGC 2415
Db 2221 ACTTTGGGAGCCGAGGACAGTGGATCACCTGAGTCAAGGATTCAGAGCTAGCTGCGC 2280
Qy 2416 AACATCATGAAACCCCTCTCTACTTAAATAATACAAAATTTAGCCAGGCGATGGTGGCAC 2475
Db 2281 AACATCATGAAACCCCTCTCTACTTAAATAATACAAAATTTAGCCAGGCGATGGTGGCAC 2340

| | | | |
|----|------|--|------|
| Qy | 2476 | ACCTGTAGTCCAGCTACTCTGGGAGGCCAACGACGACGAGAACCGCTTTGTAACCCAGAGAGTG | 2535 |
| Db | 2341 | ACCTGTAGTCCAGCTACTCTGGGAGGCCAACGACGAGAACCGCTTTGTAACCCAGAGAGTG | 2400 |
| Qy | 2536 | GAGGTTGCAGTGAGCCGAGTTTGACCAATTCGACTCCACCCCTGGCGACAGACGACGATT | 2595 |
| Db | 2401 | GAGGTTGCAGTGAGCCGAGTTTGACCAATTCGACTCCACCCCTGGCGACGACGACGATT | 2460 |
| Qy | 2596 | CTATCAAAAAAAGGACGCTGGCAAGTAAGTTATAGAAGAGAAATCTCTCTAGAGGAA | 2655 |
| Db | 2461 | CTATCAAAAAAAGGACGCTGGCAAGTAAGTTATAGAAGAGAAATCTCTCTAGAGGAA | 2520 |
| Qy | 2656 | TTAAGCGTTGTAGTAAACGGGTGCTCATCTCTTAAGCTTCGAGAAGGAGACGAAATCC | 2715 |
| Db | 2521 | TTAAGCGTTGTAGTAAACGGGTGCTCATCTCTTAAGCTTCGAGAAGGAGACGAAATCC | 2580 |
| Qy | 2716 | ATTGTGTTTAAATTCACATCTCAAGGAGGGAGAACCCGGGCTGTGTTGGGTGGTTGCCAAT | 2775 |
| Db | 2581 | ATTGTGTTTAAATTCACATCTCAAGGAGGGAGAACCCGGGCTGTGTTGGGTGGTTGCCAAT | 2640 |
| Qy | 2776 | TTCCCTAGAACGGAAATGTGGGGTATAGAAAAAGGAATGAAATAGCGTTGTTTTTCAAAT | 2835 |
| Db | 2641 | TTCCCTAGAACGGAAATGTGGGGTATAGAAAAAGGAATGAAATAGCGTTGTTTTTCAAAT | 2700 |
| Qy | 2836 | AGGTCCTCTGTAGTTATTTGATAGAGGGAAAAAGATTGACTGGGGAGGGCTTAAAAATGAT | 2895 |
| Db | 2701 | AGGTCCTCTGTAGTTATTTGATAGAGGGAAAAAGATTGACTGGGGAGGGCTTAAAAATGAT | 2760 |
| Qy | 2896 | TTGGGAAAAACAATTGCTTTTTGAGGCTCAGTGA CAACGGCAAGATTACAACCTT | 2948 |
| Db | 2761 | TTGGGAAAAACAATTGCTTTTTGAGGCTCAGTGA CAACGGCAAGATTACAACCTT | 2813 |

RESULT 5

```

US-10-342-887-937
; Sequence 937, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 937
; LENGTH: 2813
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-937

```

| | Query Match | 91.7% | Score 2813; | DB 18; | Length 2813; |
|----|-----------------------|--|---------------|-----------|--------------|
| | Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| | Matches 2813; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 136 | CTCCACCGACAGCCAGGCACTGGGCAGCAGCAGCACTGGAGACCCAGGACCCTGTGCAGGAG | 195 | | |
| Db | 1 | CTCCACCGACAGCCAGGCACTGGGCAGCAGCAGCACTGGAGACCCAGGACCCTGTGCAGGAG | 60 | | |
| Qy | 196 | CAGCTCCGGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCAGGAGCAATG | 255 | | |

QY 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1395
DB 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1260
QY 1396 GTTAAAGAGGACTCAGTCCCAACAGTGGAGGAGAAATGTTGTTGAGTCAACCACTA 1455
DB 1261 GTTAAAGAGGACTCAGTCCCAACAGTGGAGGAGAAATGTTGTTGAGTCAACCACTA 1320
QY 1456 GAGATTATAAGTCCAAAGAAAGTAGAATCAGCTTTCAAAACAGTGGACCTCAACGAAGGA 1515
DB 1321 GAGATTATAAGTCCAAAGAAAGTAGAATCAGCTTTCAAAACAGTGGACCTCAACGAAGGA 1380
QY 1516 GATGCTGACCTGAACCCACAGAACCGGAAATCTCAAAAGAGAAAGCAAAACCAAGAAC 1575
DB 1381 GATGCTGACCTGAACCCACAGAACCGGAAATCTCAAAAGAGAAAGCAAAACCAAGAAC 1440
QY 1576 TCTCTGATGGCTTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGGATCAACCCACTCA 1635
DB 1441 TCTCTGATGGCTTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGGATCAACCCACTCA 1500
QY 1636 GAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGACTATC 1695
DB 1501 GAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGACTATC 1560
QY 1696 ACACCGCCAGAGCTGAACCAACAGAGGACCAACAGAGGGTAAAGAGGGCTCTCGAAG 1755
DB 1561 ACACCGCCAGAGCTGAACCAACAGAGGACCAACAGAGGGTAAAGAGGGCTCTCGAAG 1620
QY 1756 GACAGAAGTCAAGCAGCGGAGATGAACAGCAGAGAGCAACAGCAGGAAGCCAAAGAA 1815
DB 1621 GACAGAAGTCAAGCAGCGGAGATGAACAGCAGAGAGCAACAGCAGGAAGCCAAAGAA 1680
QY 1816 CCAGCCAGTGCACAGAGCAGGCCACGGTGGACACGAACTCACTCAGAAATGGGGACAAG 1875
DB 1681 CCAGCCAGTGCACAGAGCAGGCCACGGTGGACACGAACTCACTCAGAAATGGGGACAAG 1740
QY 1876 CTCGAAAAGAGACTGAGAAGCGGACAGTCCCTTGGGGGCTCTTTAAAGGCTGGGA 1935
DB 1741 CTCGAAAAGAGACTGAGAAGCGGACAGTCCCTTGGGGGCTCTTTAAAGGCTGGGA 1800
QY 1936 CCNAGCGGATGTTGGATGCTCAAGTGCNAAACAGACCCAGTATCCATCGGCCAGTGGC 1995
DB 1801 CCNAGCGGATGTTGGATGCTCAAGTGCNAAACAGACCCAGTATCCATCGGCCAGTGGC 1860
QY 1996 AAACCCAAAGTAAACAAATCAGCACCGGTTCCACACAGGTTCTCTGCCACCAAGATGTT 2055
DB 1861 AAACCCAAAGTAAACAAATCAGCACCGGTTCCACACAGGTTCTCTGCCACCAAGATGTT 1920
QY 2056 CTCCTTACTCCATCTCTCCCAACACAGCTCCATGATATATTTCTGATGGCCAGCA 2115
DB 1921 CTCCTTACTCCATCTCTCCCAACACAGCTCCATGATATATTTCTGATGGCCAGCA 1980
QY 2116 AATGAAATTCGCTAGAAATTAAGCCGAGCTGTTGATATGAGGTGATATTTTACG 2175
DB 1981 AATGAAATTCGCTAGAAATTAAGCCGAGCTGTTGATATGAGGTGATATTTTACG 2040
QY 2176 TCTCTGCTCCAGTCTTTTCTGCAAAATAACAGTAAAGATGGTTTACGAGGTCACTAGTT 2235
DB 2041 TCTCTGCTCCAGTCTTTTCTGCAAAATAACAGTAAAGATGGTTTACGAGGTCACTAGTT 2100
QY 2236 GGGTCAGAAGAGTGCATGATCAACAGCAGGAAGGGAGGAATAGAGGAATGTTTCGG 2295
DB 2101 GGGTCAGAAGAGTGCATGATCAACAGCAGGAAGGGAGGAATAGAGGAATGTTTCGG 2160
QY 2296 GTTAAGTATGAAATGAGGTGGTGGCGGGGCTGGTGGCTCTCGCTCTAAATCTCAGC 2355
DB 2161 GTTAAGTATGAAATGAGGTGGTGGCGGGGCTGGTGGCTCTCGCTCTAAATCTCAGC 2220
QY 2356 ACTTTGGAGGCCGAGGAGGTGGATCACTGAGGTCAAGGATTCAGAGCTAGCCTGGCC 2415
DB 2221 ACTTTGGAGGCCGAGGAGGTGGATCACTGAGGTCAAGGATTCAGAGCTAGCCTGGCC 2280

QY 2416 AACATCATGAACCCCGTCTCTACTTAAATAATACAAAAATTAGCCAGGCATGGTGCACAC 2475
DB 2281 AACATCATGAACCCCGTCTCTACTTAAATAATACAAAAATTAGCCAGGCATGGTGCACAC 2340
QY 2476 ACCTGTAGTCCAGTACTCGGGAGCCCAACGACGAGAACCGTTGTACCCAGGAGTG 2535
DB 2341 ACCTGTAGTCCAGTACTCGGGAGCCCAACGACGAGAACCGTTGTACCCAGGAGTG 2400
QY 2536 GAGTTGAGTGAAGCCGAAAGTTGCACCATTCACCTCCACCTGGGCGACAGCAAGATT 2595
DB 2401 GAGTTGAGTGAAGCCGAAAGTTGCACCATTCACCTCCACCTGGGCGACAGCAAGATT 2460
QY 2596 CTATCAAAAAAAGAGCAGTGGCAAGTAAAGTTATAGAGAAATGCTCTAGNAGAA 2655
DB 2461 CTATCAAAAAAAGAGCAGTGGCAAGTAAAGTTATAGAGAAATGCTCTAGNAGAA 2520
QY 2656 TTAAGCGTTGTAGTAAACCGCTGCTCATCTCTAAGCTTGAAGAGGGAGACGAAATCC 2715
DB 2521 TTAAGCGTTGTAGTAAACCGCTGCTCATCTCTAAGCTTGAAGAGGGAGACGAAATCC 2580
QY 2716 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCCGGCTGTGTGGTGGTGGTCCAAAT 2775
DB 2581 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCCGGCTGTGTGGTGGTGGTCCAAAT 2640
QY 2776 TTCTAGNACCGAATGTTGGGTTATAGAAAAAGAAATGAATAGCTGTTTCAAT 2835
DB 2641 TTCTAGNACCGAATGTTGGGTTATAGAAAAAGAAATGAATAGCTGTTTCAAT 2700
QY 2836 AGGTCTCTGTAAAGTTATTGATGAGAGGAAAGATTCAGCTGGGAGGCTTAAATGAT 2895
DB 2701 AGGTCTCTGTAAAGTTATTGATGAGAGGAAAGATTCAGCTGGGAGGCTTAAATGAT 2760
QY 2896 TTGGAAAAACAATTCCTTTTGGGCTCAGTGACAAACCGCAAGATTTCAACTT 2948
DB 2761 TTGGAAAAACAATTCCTTTTGGGCTCAGTGACAAACCGCAAGATTTCAACTT 2813

RESULT 6

US-10-723-860-1812
; Sequence 1812, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1812
; LENGTH: 2813
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-1812

Query Match 91.7%; Score 2813; DB 20; Length 2813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 CTCACACGACGACGACGACCTGGGACGACCTGGAGACCAGGACCTGTGCAGAG 195
DB 1 CTCACACGACGACGACGACCTGGGACGACCTGGAGACCAGGACCTGTGCAGAG 60
QY 196 CAGTCCCGGTGACACGAGGGGACTGAAGATCTCCACAGGGGCTCAGCAGGAGCAATG 255
DB 61 CAGTCCCGGTGACACGAGGGGACTGAAGATCTCCACAGGGGCTCAGCAGGAGCAATG 120
QY 256 GGTAACCAAAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGATGAACCAAGAGCAGAG 315

Db 121 GGTAAACCAATGAGTGTCCCAAGAGTTGAAGACCAAGAGAAATGACCAAGAGCAGAG 180
Qy 316 ACTTACCAGGACAAACCGGTCTGCTGAACGGGGTCCAGTGGTGGTGTGCTGCACCCACACA 375
Db 181 ACTTACCAGGACAAACCGGTCTGCTGAACGGGGTCCAGTGGTGGTGTGCTGCACCCACACA 240
Qy 376 GTTCAGCACTTAGAGGAAGTCGACTTGGGAATAAGTGTCAAGACGGATAAATGTGGCCACT 435
Db 241 GTTCAGCACTTAGAGGAAGTCGACTTGGGAATAAGTGTCAAGACGGATAAATGTGGCCACT 300
Qy 436 TCTTCCCCGAGACAAACCGGAGATAGTCTGTGGGATGCGATGCCAAGGAAAGATTTGGG 495
Db 301 TCTTCCCCGAGACAAACCGGAGATAGTCTGTGGGATGCGATGCCAAGGAAAGATTTGGG 360
Qy 496 AAAGAGGCCAAACCGGAGGACCAAGCTGCTAAATCTCGTTTCTTGTGATGCTCTTCGG 555
Db 361 AAAGAGGCCAAACCGGAGGACCAAGCTGCTAAATCTCGTTTCTTGTGATGCTCTTCGG 420
Qy 556 CTTGTACAGGACGTACCGGAGACCAAGCCGAGATTCAATCCCTTGGATCAGTGAAGCTT 615
Db 421 CTTGTACAGGACGTACCGGAGACCAAGCCGAGATTCAATCCCTTGGATCAGTGAAGCTT 480
Qy 616 GATGTCACTTCAATAAAGCTCCAGCGAAACAAAGACCCAAAGTGAAGCTGACACTTCCG 675
Db 481 GATGTCACTTCAATAAAGCTCCAGCGAAACAAAGACCCAAAGTGAAGCTGACACTTCCG 540
Qy 676 GTGCGAGCTGACCGGGGACGACACAGATAAACCCAGGGGACGCCCCGGCCCAAGAC 735
Db 541 GTGCGAGCTGACCGGGGACGACACAGATAAACCCAGGGGACGCCCCGGCCCAAGAC 600
Qy 736 AAGGTCTCTCTGCGCGCAGGGATCCACAGCTTCTCCACCTGACACAGGGGGAGCAGGA 795
Db 601 AAGGTCTCTCTGCGCGCAGGGATCCACAGCTTCTCCACCTGACACAGGGGGAGCAGGA 660
Qy 796 GGAGAACTCTCCCAAGCCCAAGAGCTCCAGCTTTTGTGACAAATTTTCAAGCTGGAC 855
Db 661 GGAGAACTCTCCCAAGCCCAAGAGCTCCAGCTTTTGTGACAAATTTTCAAGCTGGAC 720
Qy 856 AAGGACAGGAAAGGTGCCAGGTGCACAGCCAAACAGGAAGCCCAAGGGGACAGCATCAA 915
Db 721 AAGGACAGGAAAGGTGCCAGGTGCACAGCCAAACAGGAAGCCCAAGGGGACAGCATCAA 780
Qy 916 GACAAAGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCTCCAGGGGAG 975
Db 781 GACAAAGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCTCCAGGGGAG 840
Qy 976 GACATAGTTGACGCAAGGAAAGAGAGACAAAGAACTTGGAACTGCGGGATGCTCTGTC 1035
Db 841 GACATAGTTGACGCAAGGAAAGAGAGACAAAGAACTTGGAACTGCGGGATGCTCTGTC 900
Qy 1036 CTTGGGACCCAGAGGACTGGAGCTGCNAAGGACGATCCCGAGGACGAGCTATAGCA 1095
Db 901 CTTGGGACCCAGAGGACTGGAGCTGCNAAGGACGATCCCGAGGACGAGCTATAGCA 960
Qy 1096 GAGATAATAATTCATCATGAGTTCTTTTAAACTCTGGTTTCACTTAAACAAAGCTGAA 1155
Db 961 GAGATAATAATTCATCATGAGTTCTTTTAAACTCTGGTTTCACTTAAACAAAGCTGAA 1020
Qy 1156 ACAAAAAAGGCCCAAGAGACACGGGTGCTGAAAAAGTCAACCACTTCAAGCTGACCTT 1215
Db 1021 ACAAAAAAGGCCCAAGAGACACGGGTGCTGAAAAAGTCAACCACTTCAAGCTGACCTT 1080
Qy 1216 AAGTCAGACAAAGCCAACTTTTATCATCCAGGAGACCCAGGGGCTGGCAAGATTTCCAAA 1275
Db 1081 AAGTCAGACAAAGCCAACTTTTATCATCCAGGAGACCCAGGGGCTGGCAAGATTTCCAAA 1140
Qy 1276 GGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTGAACTCCGAAAGGAGGC 1335
Db 1141 GGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTGAACTCCGAAAGGAGGC 1200
Qy 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCGCTCTGGGCAAACTGTTTTGGAAAAAGTCA 1395

Db 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAAGTCA 1260
Qy 1396 GTTAAAGAGGACTCAGTCCCCACAGAGTGGAGAGAAATGTGGTGTGAGTCAACCAAGTA 1455
Db 1261 GTTAAAGAGGACTCAGTCCCCACAGAGTGGAGAGAAATGTGGTGTGAGTCAACCAAGTA 1320
Qy 1456 GAGATTATAAGTCCAAAGGAGTAGAATCAGCTTCAAAAAGTGGGACCTTCAACGAAGGA 1515
Db 1321 GAGATTATAAGTCCAAAGGAGTAGAATCAGCTTCAAAAAGTGGGACCTTCAACGAAGGA 1380
Qy 1516 GATGCTCAGCTGACCCACAGAGCGAAACTCAAAAGAGAGAAACCAACCAAGGAACC 1575
Db 1381 GATGCTCAGCTGACCCACAGAGCGAAACTCAAAAGAGAGAAACCAACCAAGGAACC 1440
Qy 1576 TCTCTGATGGGCTTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCACCCACTCA 1635
Db 1441 TCTCTGATGGGCTTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCACCCACTCA 1500
Qy 1636 GAAGAAATAATGGGAAGACTCCAGCTGCGAAACATCAGACTCCAAGAAAAAGCTATC 1695
Db 1501 GAAGAAATAATGGGAAGACTCCAGCTGCGAAACATCAGACTCCAAGAAAAAGCTATC 1560
Qy 1696 ACACCGCAGAGCTGAAACCAAGAGGACCAACAGAGGGTAAAGGGCTCTCGAAG 1755
Db 1561 ACACCGCAGAGCTGAAACCAAGAGGACCAACAGAGGGTAAAGGGCTCTCGAAG 1620
Qy 1756 GACAAGAGTCAAGCAGCGGAGATGAAACAGAGAGAGCAACAGAGGAGAAAGCAAGAA 1815
Db 1621 GACAAGAGTCAAGCAGCGGAGATGAAACAGAGAGAGCAACAGAGGAGAAAGCAAGAA 1680
Qy 1816 CCAGCCAGTGCAACAGAGCAGGCGCACCGTGGACACGAACTCCTCAGCAAGATGGGGACAAG 1875
Db 1681 CCAGCCAGTGCAACAGAGCAGGCGCACCGTGGACACGAACTCCTCAGCAAGATGGGGACAAG 1740
Qy 1876 CTCCAAAGAGACCTGAGAGCGGCGCAGCTCCCTTGGGGCTCTTTTAAAGGCTGGGA 1935
Db 1741 CTCCAAAGAGACCTGAGAGCGGCGCAGCTCCCTTGGGGCTCTTTTAAAGGCTGGGA 1800
Qy 1936 CCAAAGCGGATGTTGGATGCTCAAAGTCAAAACAGAGCCAGTATCCATCGGACCAAGTTGGC 1995
Db 1801 CCAAAGCGGATGTTGGATGCTCAAAGTCAAAACAGAGCCAGTATCCATCGGACCAAGTTGGC 1860
Qy 1996 AAACCCAGTAAACAAATCAGCAAGTTCACAGGTTCTCCTGCGCAACCAAGATGTTT 2055
Db 1861 AAACCCAGTAAACAAATCAGCAAGTTCACAGGTTCTCCTGCGCAACCAAGATGTTT 1920
Qy 2056 CTCTCTACTCTCTCTCCCAACAGCTCCATGTATATATTTCTCTGATGSCCAGCA 2115
Db 1921 CTCTCTACTCTCTCTCCCAACAGCTCCATGTATATATTTCTCTGATGSCCAGCA 1980
Qy 2116 AATGAAATTCCTGCTAGAAATTAAGCCCGAGCTGTTGTATATTGAGGTGTATTATTACG 2175
Db 1981 AATGAAATTCCTGCTAGAAATTAAGCCCGAGCTGTTGTATATTGAGGTGTATTATTACG 2040
Qy 2176 TCTCTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTGTAGCAGGTCACTAGTT 2235
Db 2041 TCTCTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTGTAGCAGGTCACTAGTT 2100
Qy 2236 GGGTCAAGAGTCTGATGATCAACCAAGAGGAGGAGGGAATAGAGGAATGTTTCGG 2295
Db 2101 GGGTCAAGAGTCTGATGATCAACCAAGAGGAGGAGGGAATAGAGGAATGTTTCGG 2160
Qy 2296 GTTAAAGTATGAAATGGCAGTGGTGGCGGGCGTGGTGGCTCTCGCTCTGAATCTCAGC 2355
Db 2161 GTTAAAGTATGAAATGGCAGTGGTGGCGGGCGTGGTGGCTCTCGCTCTGAATCTCAGC 2220
Qy 2356 ACTTTGGAGGCCGAGGAGGTGATCACTGAGGTCAAGGATTCAGAGCTCAAGCTAGCTGGCC 2415
Db 2221 ACTTTGGAGGCCGAGGAGGTGATCACTGAGGTCAAGGATTCAGAGCTCAAGCTAGCTGGCC 2280
Qy 2416 AACATCATGAAACCCCGTCTCTCTTAAATAATTAAGGAGGATGAGTGGCACAC 2475
Db 2281 AACATCATGAAACCCCGTCTCTCTTAAATAATTAAGGAGGATGAGTGGCACAC 2340

Qy 907 GAGCATCAAGACAAAGGTGGATGAGTTCTCTGGCTTTATCAGGCGAGTCGATGATGTCCCT 966
Db |||||||
Qy 786 GAGCATCAAGACAAAGGTGGATGAGTTCTCTGGCTTTATCAGGCGAGTCGATGATGTCCCT 845
Db |||||||
Qy 967 GCAGGGAAGGACATAGTTGACGGCAAGGAAAGAAAGAGCAAGAACTTGGAACTGCGGAT 1026
Db |||||||
Qy 846 GCAGGGAAGGACATAGTTGACGGCAAGGAAAGAAAGAGCAAGAACTTGGAACTGCGGAT 905
Db |||||||
Qy 1027 TGCTCTGTCCCTGGGGACCCAGAGGACTGGAGACTGCAAGAGGAGTTCACCGGACGCA 1086
Db |||||||
Qy 906 TGCTCTGTCCCTGGGGACCCAGAGGACTGGAGACTGCAAGAGGAGTTCACCGGACGCA 965
Db |||||||
Qy 1087 GCTATAGCAGAGAAATAATTTCCATCATGAGTTTCTTTAAATCTCTGGTTTCACTTAAC 1146
Db |||||||
Qy 966 GCTATAGCAGAGAAATAATTTCCATCATGAGTTTCTTTAAATCTCTGGTTTCACTTAAC 1025
Db |||||||
Qy 1147 AAGCTGAAACAAAAGGACCCAGAGACACAGCGGTGCTGAAAAGTCAACCACTTCA 1206
Db |||||||
Qy 1026 AAGCTGAAACAAAAGGACCCAGAGACACAGCGGTGCTGAAAAGTCAACCACTTCA 1085
Db |||||||
Qy 1207 GCTGACCTTAAGTCAGACAAAGCCAACTTTATATCCAGGAGACCCAGGGGCTGGCAAG 1266
Db |||||||
Qy 1086 GCTGACCTTAAGTCAGACAAAGCCAACTTTATATCCAGGAGACCCAGGGGCTGGCAAG 1145
Db |||||||
Qy 1267 AATTCCAAAGGATGCAACCCATCGGGGCACACACAGTCCGTGACAAACCCCTGAACCTGCG 1326
Db |||||||
Qy 1146 AATTCCAAAGGATGCAACCCATCGGGGCACACACAGTCCGTGACAAACCCCTGAACCTGCG 1205
Db |||||||
Qy 1327 AAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTCTGGCTCTGGGCAAACTGTGTTGG 1386
Db |||||||
Qy 1206 AAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTCTGGCTCTGGGCAAACTGTGTTGG 1265
Db |||||||
Qy 1387 AAAAAGTCAGTTAAAGAGACTCAGTCCCAAGTGGGAGGAGAAATGTGTGTGAG 1446
Db |||||||
Qy 1266 AAAAAGTCAGTTAAAGAGACTCAGTCCCAAGTGGGAGGAGAAATGTGTGTGAG 1325
Db |||||||
Qy 1447 TCACCACTAGAGATTATAAGTCCAAAGGAAGTAGAATCAGCTTTACAAACAGTGGACCTC 1506
Db |||||||
Qy 1326 TCACCACTAGAGATTATAAGTCCAAAGGAAGTAGAATCAGCTTTACAAACAGTGGACCTC 1385
Db |||||||
Qy 1507 AACGAAGGAGATGTGCACTGAAACCCACAGAGGAAACTCAAAAGAGAGAAAGCAAA 1566
Db |||||||
Qy 1386 AACGAAGGAGATGTGCTCTGCTGAAACCCACAGAGGAAACTCAAAAGAGAGAAAGCAAA 1445
Db |||||||
Qy 1567 CCAAGAACTCTCTGATGGCGTTTCTCAGACAAATGTCTAGTGAAGGGGATGAGGGATC 1626
Db |||||||
Qy 1446 CCAAGAACTCTCTGATGGCGTTTCTCAGACAAATGTCTCAGACAAATGTCTCAGACAAATG 1481
Db |||||||
Qy 1627 ACCCACTCAGAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAA 1686
Db |||||||
Qy 1482 -----ACATCAGACTCCACAGAA 1499
Db |||||||
Qy 1687 AAGACTATCACCCGACAGGCTGAAACCAAGAGGACCAAGAGGCTGAAAGAGGGC 1746
Db |||||||
Qy 1500 AAGACTATCACCCGACAGGCTGAAACCAAGAGGACCAAGAGGCTGAAAGAGGGC 1559
Db |||||||
Qy 1747 TCCTCGAAGGACAAAGTCTCAGCAGCGAGATGAACAGCAGAGCAACAGCAGGAA 1806
Db |||||||
Qy 1560 TCCTCGAAGGACAAAGTCTCAGCAGCGGAGATGAACAGCAGAGCAACAGCAGGAA 1619
Db |||||||
Qy 1807 GCCAAGAACCCAGCCAGTGCAACAGAGGCGCCAGGTGGACACGAACTCAGTGCAGAAAT 1866
Db |||||||
Qy 1620 GCCAAGAACCCAGCCAGTGCAACAGAGGCGCCAGGTGGACACGAACTCAGTGCAGAAAT 1679
Db |||||||
Qy 1867 GGGGCAAGACTCAAAAGAGACTCAGAGAGCGGACAGTCCCTTGGGGGCTTTCTTTAAA 1926
Db |||||||
Qy 1680 GGGGCAAGACTCAAAAGAGACTCAGAGAGCGGACAGTCCCTTGGGGGCTTTCTTTAAA 1739
Db |||||||
Qy 1927 GGCCTGGGACCAAGCGGATGTTGATGCTCAAGTGCAGAGCCAGACCCAGTATCCATCGGA 1986
Db |||||||
Qy 1740 GGCCTGGGACCAAGCGGATGTTGATGCTCAAGTGCAGAGCCAGACCCAGTATCCATCGGA 1799
Db |||||||

Qy 1987 CCAAGTTGGCAAAACCAAGTAATAACAAATCAGACGGTTCCCAACCAAGTTCTCTGCCACCA 2046
Db |||||||
Qy 1800 CCAAGTTGGCAAAACCAAGTAATAACAAATCAGACGGTTCCCAACCAAGTTCTCTGCCACCA 1859
Db |||||||
Qy 2047 AGATGTGTTCTCTTACTCCATCTCTCTCCCAACACGCTCCATGTATATATCTTCTGA 2106
Db |||||||
Qy 1860 AGATGTGTTCTCTTACTCCATCTCTCTCCCAACACGCTCCATGTATATATCTTCTGA 1919
Db |||||||
Qy 2107 TGGCCAGCAAAATGAAATTTCTGCCCTAGAAATTAAGCCCGAGCTGTGTATATTTGAGGTGTA 2166
Db |||||||
Qy 1920 TGGCCAGCAAAATGAAATTTCTGCCCTAGAAATTAAGCCCGAGCTGTGTATATTTGAGGTGTA 1979
Db |||||||
Qy 2167 TTATTTACGCTCTCTGGTCCAGTCTTTTCTGCCAAATAACAG 2207
Db |||||||
Qy 1980 TTATTTACGCTCTCTGGTCCAGTCTTTTCTGCCAAATAACAG 2020
Db |||||||
RESULT 8
US-09-939-825-15
; Sequence 15, Application US/09939825
; Patent No. US20020151009A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: PT046P1
; FILE REFERENCE: Human Polynucleotides, Polypeptides, and Antibodies
; CURRENT APPLICATION NUMBER: US/09/939,825
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: PCT/US01/05498
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/184,664
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/189,874
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-939-825-15
Query Match 51.1%; Score 1567; DB 9; Length 2105;
Best Local Similarity 85.1%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 190; Indels 142; Gaps 7;
Qy 744 CTCTGCCGCCAGGAGTCCACGCTTCTCCACCTGAGACAGGGGAGCAGCAGAGAGAGC 803
Db |||||||
Qy 9 CTCTGCCGCCAGGAGTCCACGCTTCTCCACCTGAGACAGGGGAGCAGGAGAGAGC 68
Db |||||||
Qy 804 TCCCTCCAAGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTCAAGCTGGACAGGGACA 863
Db |||||||
Qy 69 TCCCTCCAAGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTCAAGCTGGACAGGGACA 128
Db |||||||
Qy 864 GGAAGAGTCCAGGTGACAGCCAAACAGGAGCCAAAGGGCAGAGCATCAAGCAAGT 923
Db |||||||
Qy 129 GGAAGAGTCCAGGTGACAGCCAAACAGGAGCCAAAGGGCAGAGCATCAAGCAAGT 188
Db |||||||
Qy 924 GGATGAGTTCCTCGCTTATCAGGGCAGTCCGATGATGTCCTCGCAGGAAAGGACATAGT 983
Db |||||||
Qy 189 GGATGAGTTCCTCGCTTATCAGGGCAGTCCGATGATGTCCTCGCAGGAAAGGACATAGT 248
Db |||||||
Qy 984 TGACGGCAAGGAAAAAGAGCAAGAACTTGGAACTGCGGATTCCTGCTCCCTGGGGA 1043
Db |||||||
Qy 249 TGACGGCAAGGAAAAAGAGCAAGAACTTGGAACTGCGGATTCCTGCTCCCTGGGGA 308
Db |||||||
Qy 1044 CCCAGAGGACTGGAGACTGCAAGAGCAGTCCAGGACAGCATATAGCAGAGATAA 1103
Db |||||||
Qy 309 CCCAGAGGACTGGAGACTGCAAGAGCAGTCCAGGACAGCATATAGCAGAGATAA 368
Db |||||||
Qy 1104 TAAATCCATCATGAGTTTCTTTTAAATCTCTGGTTTACCTAACAAAGCTGAAAAA 1163
Db |||||||
Qy 369 TAAATCCATCATGAGTTTCTTTTAAATCTCTGGTTTACCTAACAAAGCTGAAAAA 428
Db |||||||

1164 QY GGACCCAGAGACACGGGTGCTGAAAAGTCACCCACCACTTCAGCTGACCTTTAAGTCAGA 1223
429 Db GGACCCAGAGACACGGGATCGAAA-----CGAGAAAGTGCTCT 466
1224 QY CAAAGCCAACTTTTATCATCCAGGAGACCCAGGGCTGGCAAGAAATCCAAAGAGATGCAA 1283
467 Db GTGATGGCAACAGCTGGTCAGAGACATCCGAGATCCAGGCTAGAGGCAACCAAGAA--AA 523
1284 QY CCCATCGGGGCACACAGATCGGTGCTGACCAACCCCTGAACTCGGAAGGAGGCACCAAGGA 1343
524 Db AGCAGCTGGATAGCCCAAGGCTAGGACTCGCCCTTTA----- 559
1344 QY GAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAAGTCAGTTAAAGA 1403
560 Db -----GAAATTTCTTTAGGCAATAGGGTGTCTGAAAA 590
1404 QY GGACTCAGTCCCAACAGGTGGGAGGAGAAATGTGTGTGTGAGTCACCAAGTAGAGATTAT 1463
591 Db GTCAACCCACCACTTCAGCTGACCTTAAAG-----TCAGACAAAGCCAACTTTACATCCCGAG 645
1464 QY AAAGTCCAAAGAGTAGAATGAGCTTACAAACAGTGGACCTCAACGAAAGGAGATGCTGC 1523
646 Db GAGACCCAAAGGGCTGGCA--AGAAATCCAAAGGATGCAACCCATCGGGGCACACACAGT 703
1524 QY ACCTGAACCCACAGAAAGCGAAACTCAAAGAGAGAAAGCAACCAAGAACTCTCTGAT 1583
704 Db CGGTGACAAACCCCTGAACTCTGGAGGAGGACCAAGGAGAAATCAGGACCCACCTCTC 763
1584 QY GCGCTTTCTCAGACAAATGTCTAGTGAAGGGGATGGAGGGATCACCCACTCAGAGAAAT 1643
764 Db TGCCTCT----- 770
1644 QY AAATGGGAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGATATCACAACGCC 1703
771 Db ---GGGCAAACTGTTTTGGAAAAAGACATCAGACTCCACAGAAAAAGACTATCACAACGCC 827
1704 QY AGAGCTCAACCCACAGAGCACCACAGAGGTTAAAGGGCTCCTCGAAGGCAAGAA 1763
828 Db AGAGCTCAACCCACAGAGCACCACAGAGGTTAAAGGGCTCCTCGAAGGCAAGAA 887
1764 QY CTCAGCAGCCGAGATGAACAGCAGAGCAACAGCAGGAAGCAACAGCAACAGCCCA 1823
888 Db CTCAGCAGCCGAGATGAACAGCAGAGCAACAGCAGGAAGCAACAGCAACAGCCCA 947
1824 QY GTGCACAGAGAGGCCACGGTGGACAGAACTCACTGAGAAATGGGCAAGCTCCAAAA 1883
948 Db GTGCACAGAGAGGCCACGGTGGACAGAACTCACTGAGAAATGGGCAAGCTCCAAAA 1007
1884 QY GAGACCTCAGAGCGGCAGCAGTCCCTTGGGGGCTTCTTTAAGGCTGGGACCAAGCG 1943
1008 Db GAGACCTCAGAGCGGCAGCAGTCCCTTGGGGGCTTCTTTAAGGCTGGGACCAAGCG 1067
1944 QY GATGTTGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGACAGTTGGCAACCCAA 2003
1068 Db GATGTTGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGACAGTTGGCAATCCAA 1127
2004 QY GTAAACAAATCAGCAGGTTCCCAACAGGTTCTCTGCAACCAAGATGTCTCTCTTAC 2063
1128 Db GTAAACAAATCAGCAGGTTCCCAACAGGTTCTCTGCAACCAAGATGTCTCTCTTAC 1187
2064 QY TCCATCTCTCCCAACAGCAGTCAATATATTTCTTCTGATGGCCAGCAAGTAAAT 2123
1188 Db TCCATCTCTCCCAACAGCAGTCAATATATTTCTTCTGATGGCCAGCAAGTAAAT 1247
2124 QY TCTGCTAGAAATTAAGCCGAGCTGTGTATATTTAGGCTGTATTTATGCTCTCTGGT 2183
1248 Db TCTGCTAGAAATTAAGCCGAGCTGTGTATATTTAGGCTGTATTTATGCTCTCTGGT 1307
2184 QY CCAGTCTTTTCTGGCAAAATACAGTAAGATGTTTACAGGTCACCTAGTTGGGTGAGA 2243
1308 Db CCAGTCTTTTCTGGCAAAATACAGTAAGATGTTTACAGGTCACCTAGTTGGGTGAGA 1367
2244 QY AGAGTCGATGATCACCAGCAGGAAAGGGAGGGAATAGAGGAATGTGTTCGGGTTAAGTG 2303

1368 Db AGAGTCGATGATCACCAGCAGGAAAGGAGGGAATAGAGGAATGTGTTCGGGTTAAGTG 1427
2304 QY ATGAAAAATGGCAGTGGTGGCCGGCGTGGCTCTCGCCTGTATATCTCAGCACTTTGGG 2363
1428 Db ATGAAAAATGGCAGTGGTGGCCGGCGTGGCTCTCGCCTGTATATCTCAGCACTTTGGG 1487
2364 QY AGGCGGAGGAGGTGGATCAGCTGAGGTGAGAGTTCAAGACTAGCCTGGCCCAACATCAT 2423
1488 Db AGGCGGAGGAGGTGGATCAGCTGAGGTGAGAGTTCAAGACTAGCCTGGCCCAACATCAT 1547
2424 QY GAAACCCCGTCTCTACTAAAAATACAAAAATTTAGCCAGGATGGTGGCACAACACCTGTAG 2483
1548 Db GAAACCCCGTCTCTACTAAAAATACAAAAATTTAGCCAGGATGGTGGCACAACACCTGTAG 1607
2484 QY TCCAGCTACTCTCGGAGGCCCAACGACGAGAACCGCTTTGTATCCCAAGGAGGTGGAGTTGC 2543
1608 Db TCCAGCTACTCTCGGAGGCCCAACGACGAGAACCGCTTTGTATCCCAAGGAGGTGGAGTTGC 1667
2544 QY AGTGAGCGGAAGTTGCACCATTTGCACTCCACCTGGGGCGACAGAGCAAGATTTCTATC-AA 2602
1668 Db AGTGAGCGGAAGTTGCACCATTTGCACTCCACCTGGGGCGACAGAGCAAGATTTCTATCAA 1727
2603 QY AAAAAAGGCGAGTGGCAAGTAAGTTATAGAGAGAAATGCTGTAGAGGAATTTAAGCG 2662
1728 Db AAAAAAGGCGAGTGGCAAGTAAGTTATAGAGAGAAATGCTGTAGAGGAATTTAAGCG 1787
2663 QY TTGTAGTAAACGCGTGTCTCATCTTAAGCTTTGAAGAGGAGAGACGAAATCCCATTTGTT 2722
1788 Db TTGTAGTAAACGCGTGTCTCATCTTAAGCTTTGAAGAGGAGAGACGAAATCCCATTTGTT 1847
2723 QY TAAATTCACATCTCAAGAGAGGAGAAACCGGGCTGTGTGTGGTGTGCTCCCAATTTCTAG 2782
1848 Db TAAATTCACATCTCAAGAGAGGAGAAACCGGGCTGTGTGTGGTGTGCTCCCAATTTCTAG 1907
2783 QY AACGGAATGTGGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAATAGGGTCC 2842
1908 Db AACGGAATGTGGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAATAGGGTCC 1967
2843 QY TTGTAAGTTATTGATGAGAGGAAAAAGATTGACCTGGGAGGGCTTAAATGATTTGGGAA 2902
1968 Db TTGTAAGTTATTGATGAGAGGAAAAAGATTGACCTGGGAGGGCTTAAATGATTTGGGAA 2027
2903 QY AACAAATGCTTTTGGGCTCAGTCAACACGCAAGATTAACAATTAATAAAAAA 2962
2028 Db AACAAATGCTTTTGGGCTCAGTCAACACGCAAGATTAACAATTAATAAAAAA 2087
2963 QY AAAAA 2967
2088 Db AAAAA 2092

RESULT 9

US-10-737-082-47

; Sequence 47, Application US/10737082

; Publication No. US20050130170A1

; GENERAL INFORMATION:

; APPLICANT: Bayer Healthcare LLC

; APPLICANT: Beard, Chris

; APPLICANT: Burgess, Chris

; APPLICANT: Gannon, Allison

; APPLICANT: Harvey, Jeanne

; APPLICANT: Lechner, John F.

; APPLICANT: Li, Zheng

; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences

; FILE REFERENCE: 1857/2032

; CURRENT APPLICATION NUMBER: US/10/737,082

; CURRENT FILING DATE: 2003-12-16

; PRIOR APPLICATION NUMBER: US 10/737,082

; PRIOR FILING DATE: 2003-12-16

; NUMBER OF SEQ ID NOS: 300

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 47

; LENGTH: 127567
; TYPE: DNA
; ORGANISM: Homo. sapiens
US-10-737-082-47

Query Match 32.6%; Score 1000.6; DB 22; Length 127567;
Best Local Similarity 99.3%; Pred. No. 7.1e-269;
Matches 1016; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy 1929 CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCACACAGACCCAGTATCCATCGGACC 1988
Db 126545 CCAGGACCAAGCGGATGTTGGATGCTCAAGTGCACACAGACCCAGTATCCATCGGACC 126604

Qy 1989 AGTTGGCAAAACCCAAAGTAACAAATCAGCAGCGTTCCACCAAGGTTCTCTGCCACCAAG 2048
Db 126605 AGTTGGCAAAACCCAAAGTAACAAATCAGCAGCGTTCCACCAAGGTTCTCTGCCACCAAG 126664

Qy 2049 ATGTGTTCTCTTACTCCATCTCTCCCAACAGCTCCATGTATATATTTCTTCTGATG 2108
Db 126665 ATGTGTTCTCTTACTCCATCTCTCCCAACAGCTCCATGTATATATTTCTTCTGATG 126724

Qy 2109 GCCAGCAATGAAATTTCTGCCTAGAAATTAAGCCGAGCTGTGTATATTCAGGTGTATT 2168
Db 126725 GCCAGCAATGAAATTTCTGCCTAGAAATTAAGCCGAGCTGTGTATATTCAGGTGTATT 126784

Qy 2169 ATTTACGTTCTGTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTAGCAGGTCA 2228
Db 126785 ATTTACGTTCTGTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTAGCAGGTCA 126844

Qy 2229 CCTAGTTGGGTTCAGAGAGTCGATGATCACAAGCAGGAAAGGAGGGAATAGAGGAATG 2288
Db 126845 CCTAGTTGGGTTCAGAGAGTCGATGATCACAAGCAGGAAAGGAGGGAATAGAGGAATG 126904

Qy 2289 TGTTCGGGTTAAGTATGAAATGCAATGCACTGTGCGCGCGTGTGGCTCGCCTGTAA 2348
Db 126905 TGTTCGGGTTAAGTATGAAATGCAATGCACTGTGCGCGCGTGTGGCTCGCCTGTAA 126964

Qy 2349 TCTCAGCACTTTGGGAGCCGAGGAGGTGGATCACCTGAGGTCAAGAGTTTCAAGACTAG 2408
Db 126965 TCTCAGCACTTTGGGAGCCGAGGAGGTGGATCACCTGAGGTCAAGAGTTTCAAGACTAG 127024

Qy 2409 CCTGCCAAACATCATGAAACCCGCTCTCTAATAAAATACAAAATTAAGCAGGATGGT 2468
Db 127025 CCTGCCAAACATCATGAAACCCGCTCTCTAATAAAATACAAAATTAAGCAGGATGGT 127084

Qy 2469 GGCACACCTGTAGTCCAGCTACTCGGAGCCCAACGACGAAACCGCTGTACCCA 2528
Db 127085 GGCACACCTGTAGTCCAGCTACTCGGAGCCCAACGACGAAACCGCTGTACCCA 127144

Qy 2529 GGAGTGGAGGTTGAGTGCAGTGCAGGAGTTGCACCAATTCACCTCCACCCCTGGGCGACAG 2588
Db 127145 GGAGTGGAGGTTGAGTGCAGTGCAGGAGTTGCACCAATTCACCTCCACCCCTGGGCGACAG 127204

Qy 2589 CAAGATTTCTATC - - - AAAAAAAAAAGGAGTGGCAAGTAAGTATAGAAAGAAATGCTG 2645
Db 127205 CAAGATTTCTATCAAAAAAAAAAGGAGTGGCAAGTAAGTATAGAAAGAAATGCTG 127264

Qy 2646 CTAGAGGAATTAAGCGTTGTAGTAAACCGTGTCTATCCTCTAAGCTTGAAGAGGAG 2705
Db 127265 CTAGAGGAATTAAGCGTTGTAGTAAATGCGTGTCTATCCTCTAAGCTTGAAGAGGAG 127324

Qy 2706 ACAGAAATCCATTTGTTTAAATTCATCTCAAGGAGGAGAAACCCGGCTGTGTGGGT 2765
Db 127325 ACAGAAATCCATTTGTTTAAATTCATCTCAAGGAGGAGAAACCCGGCTGTGTGGGT 127384

Qy 2766 GGTTCGCAATTTCTTAGAACGGAATGTGTGGGTTATAGAAAAGAAATGAATAGCGTTG 2825
Db 127385 GGTTCGCAATTTCTTAGAACGGAATGTGTGGGTTATAGAAAAGAAATGAATAGCGTTG 127444

Qy 2826 TTTTTCAAATAGGGTCTTGTAAAGTATTGATGAGAGGAAAGATTGACTGGGAGGCG 2885
Db 127445 TTTTTCAAATAGGGTCTTGTAAAGTATTGATGAGAGGAAAGATTGACTGGGAGGCG 127504

Qy 2886 TTAAATGATTTGGGAAAAACAATTGCTTTTGGGCTCAGTGACACCGCAAGATTACAA 2945
Db 127505 TTAAATGATTTGGGAAAAACAATTGCTTTTGGGCTCAGTGACACCGCAAGATTACAA 127564

Qy 2946 CTT 2948
Db 127565 CTT 127567

RESULT 10
US-10-765-790-47
; Sequence 47, Application US/10765790
; Publication No. US20050130172A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 47
; LENGTH: 127567
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-765-790-47

Query Match 32.6%; Score 1000.6; DB 22; Length 127567;
Best Local Similarity 99.3%; Pred. No. 7.1e-269;
Matches 1016; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy 1929 CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCACACAGACCCAGTATCCATCGGACC 1988
Db 126545 CCAGGACCAAGCGGATGTTGGATGCTCAAGTGCACACAGACCCAGTATCCATCGGACC 126604

Qy 1989 AGTTGGCAAAACCCAAAGTAACAAATCAGCAGCGTTCCACCAAGGTTCTCTGCCACCAAG 2048
Db 126605 AGTTGGCAAAACCCAAAGTAACAAATCAGCAGCGTTCCACCAAGGTTCTCTGCCACCAAG 126664

Qy 2049 ATGTGTTCTCTTACTCCATCTCTCCCAACAGCTCCATGTATATATTTCTTCTGATG 2108
Db 126665 ATGTGTTCTCTTACTCCATCTCTCCCAACAGCTCCATGTATATATTTCTTCTGATG 126724

Qy 2109 GCCAGCAATGAAATTTCTGCCTAGAAATTAAGCCGAGCTGTGTATATTCAGGTGTATT 2168
Db 126725 GCCAGCAATGAAATTTCTGCCTAGAAATTAAGCCGAGCTGTGTATATTCAGGTGTATT 126784

Qy 2169 ATTTACGTTCTGTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTAGCAGGTCA 2228
Db 126785 ATTTACGTTCTGTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTAGCAGGTCA 126844

Qy 2229 CCTAGTTGGGTTCAGAGAGTCGATGATCACAAGCAGGAAAGGAGGGAATAGAGGAATG 2288
Db 126845 CCTAGTTGGGTTCAGAGAGTCGATGATCACAAGCAGGAAAGGAGGGAATAGAGGAATG 126904

Qy 2289 TGTTCGGGTTAAGTATGAAATGCAATGCACTGTGCGCGCGTGTGGCTCGCCTGTAA 2348
Db 126905 TGTTCGGGTTAAGTATGAAATGCAATGCACTGTGCGCGCGTGTGGCTCGCCTGTAA 126964

Qy 2349 TCTCAGCACTTTGGGAGCCGAGGAGGTGGATCACCTGAGGTCAAGAGTTTCAAGACTAG 2408
Db 126965 TCTCAGCACTTTGGGAGCCGAGGAGGTGGATCACCTGAGGTCAAGAGTTTCAAGACTAG 127024

Qy 2409 CCTGCCAAACATCATGAAACCCGCTCTCTAATAAAATACAAAATTAAGCAGGATGGT 2468
Db 127025 CCTGCCAAACATCATGAAACCCGCTCTCTAATAAAATACAAAATTAAGCAGGATGGT 127084

| | | | |
|----|--------|--|--------|
| Db | 127025 | CCTGGCCACATCATGAACCCCGTCTCTACTATAAAATATACAAAATTTAGCCAGCGCATGGT | 127084 |
| Qy | 2469 | GGCAGACACCTGTAGTCCCACTACTCGGAGCCCAACGACGACGAGAACCCGCTTGTACCCA | 2528 |
| Db | 127085 | GGCAGACACCTGTAGTCCCACTACTCGGAGCCCAACGACGACGAGAACCCGCTTGTACCCA | 127144 |
| Qy | 2529 | GGAGGTGGAGGTTGAGTGAGCCGGAAGTTGACCATTTGCACTCCACCTGGGCGACAGAG | 2588 |
| Db | 127145 | GGAGGTGGAGGTTGAGTGAGCCGGAAGTTGACCATTTGCACTCCACCTGGGCGACAGAG | 127204 |
| Qy | 2589 | CAAGATTCTATC---AAAAAAGGAGGAGGCAAGTAAAGTTATAGAGAGAAATCGTG | 2645 |
| Db | 127205 | CAAGATTCTATCAAAAAAAGGAGGAGGCAAGTAAAGTTATAGAGAGAAATCGTG | 127264 |
| Qy | 2646 | CTAGAAGGAATTAAGCGTTGTAGTAAACGCGTGTCTATCTCTTAAGCTTGAAGAAGGGAG | 2705 |
| Db | 127265 | CTAGAAGGAATTAAGCGTTGTAGTAAACGCGTGTCTATCTCTTAAGCTTGAAGAAGGGAG | 127324 |
| Qy | 2706 | ACGAAATCCATTTGTTTAAATTCACATCTCAAGAGGAGGAACCCGGGCTGTGTGGGT | 2765 |
| Db | 127325 | ACGAAATCCATTTGTTTAAATTCACATCTCAAGAGGAGGAACCCGGGCTGTGTGGGT | 127384 |
| Qy | 2766 | GGTTGCCAAATTTCTTAGAACGGAATGTGGGGTATAGAAAAAGAAATGAATAGCGTTG | 2825 |
| Db | 127385 | GGTTGCCAAATTTCTTAGAACGGAATGTGGGGTATAGAAAAAGAAATGAATAGCGTTG | 127444 |
| Qy | 2826 | TTTTTCAATAGGGTCCTTGTAAAGTTATTCATAGAGGGGAAAAAGTTGACTGGGGAGGGC | 2885 |
| Db | 127445 | TTTTTCAATAGGGTCCTTGTAAAGTTATTCATAGAGGGGAAAAAGTTGACTGGGGAGGGC | 127504 |
| Qy | 2886 | TTAAAAATGATTTGGGAAAAACAATTCGCTTTTGAGGCTCAGTGACAAACGGCAAAAGATTCAA | 2945 |
| Db | 127505 | TTAAAAATGATTTGGGAAAAACAATTCGCTTTTGAGGCTCAGTGACAAACGGCAAAAGATTCAA | 127564 |
| Qy | 2946 | CTT 2948 | |
| Db | 127565 | CTT 127567 | |

RESULT 11

US-10-029-386-22674/c

; Sequence 22674, Application US/10029386

; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: AECOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1

; SEQ ID NO 22674

; LENGTH: 644

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC004501.1

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = .2

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7

; OTHER INFORMATION: NT HIT: AF041260.1, EVALUE 0.00e+00

; OTHER INFORMATION: SWISSPROT HIT: O75363, EVALUE 6.00e-79

; OTHER INFORMATION: EST_HUMAN HIT: B1549901.1, EVALUE 0.00e+00

US-10-029-386-22674

Query Match 19.0%; Score 582.6; DB 16; Length 644;

Best Local Similarity 99.3%; Pred. No. 1.3e-152;

Matches 585; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

| | | | |
|--|-----|--|-----|
| Qy | 395 | TCGACTTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACATTTCTTCCCCCGAGAGCAAAACGG | 454 |
| Db | 644 | TCGACTTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACATTTCTTCCCCCGAGAGCAAAACGG | 585 |
| Qy | 455 | AGATAAGTGTCTGTTGGCGGATGCCCAACGGAAAGAATCTTTGGGAAAGAGGCCAAACCCGAGG | 514 |
| Db | 584 | AGATAAGTGTCTGTTGGCGGATGCCCAACGGAAAGAATCTTTGGGAAAGAGGCCAAACCCGAGG | 525 |
| Qy | 515 | CACCAGCTGTAAATCTCGTTTTTTCTTGATGCTCTCTCGGCTGTCTACAGGACGTACCG | 574 |
| Db | 524 | CACCAGCTGTAAATCTCGTTTTTTCTTGATGCTCTCTCGGCTGTCTACAGGACGTACCG | 465 |
| Qy | 575 | GAGACCAAGCCGCGAGATTCATCCCTTGGATCAGTGAAGCTTGATGTCAAGTCCAATAAAG | 634 |
| Db | 464 | GAGACCAAGCCGCGAGATTCATCCCTTGGATCAGTGAAGCTTGATGTCAAGTCCAATAAAG | 405 |
| Qy | 635 | CTCAGCGGAACAAAGACCAAGTCGAGCTGGACACTTCCGTTGGCAGCTGGACCGGGGC | 694 |
| Db | 404 | CTCAGCGGAACAAAGACCAAGTCGAGCTGGACACTTCCGTTGGCAGCTGGACCGGGGC | 345 |
| Qy | 695 | AGGACACAGATAAAACCCCGGGCAGCGCCCGGCCCAAGACAAGGTCTCTCTGCGGCCA | 754 |
| Db | 344 | AGGACACAGATAAAACCCCGGGCAGCGCCCGGCCCAAGACAAGGTCTCTCTGCGGCCA | 285 |
| Qy | 755 | GGGATCCACGCTTCTCCCACTTCAGACAGCGGGGAGCAGGAGGAAGTCTCCCTCCAAGC | 814 |
| Db | 284 | GGGATCCACGCTTCTCCCACTTCAGACAGCGGGGAGCAGGAGGAAGTCTCCCTCCAAGC | 225 |
| Qy | 815 | CAAAGGACTCAGCTTTTGTGACAAATTTCTTCAAGCTGGACAAGGGACAGGAAAAGGTGC | 874 |
| Db | 224 | CAAAGGACTCAGCTTTTGTGACAAATTTCTTCAAGCTGGACAAGGGACAGGAAAAGGTGC | 165 |
| Qy | 875 | CAGGTGACAGCCACAGAGGAGCCAAAGAGGGCAGAGCATCAAGACAAGGTGGATGAGTTTC | 934 |
| Db | 164 | CAGGTGACAGCCACAGAGGAGCCAAAGAGGGCAGAGCATCAAGACAAGGTGGATGAGTTTC | 105 |
| Qy | 935 | CTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCAGGGAAGGACATAGT | 983 |
| Db | 104 | CTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCAGGGAAGGTAAGTGT | 56 |
| RESULT 12 | | | |
| US-09-783-590-10392 | | | |
| ; Sequence 10392, Application US/09783590 | | | |
| ; Patent No. US20020110850A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: Dillon, Patrick J. | | | |
| ; APPLICANT: Haseltine, William A. | | | |
| ; APPLICANT: Li, Haodong | | | |
| ; APPLICANT: Rosen, Craig A. | | | |
| ; APPLICANT: Ruben, Steven M. | | | |
| ; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2 | | | |
| ; FILE REFERENCE: PO-16.2C1 | | | |
| ; CURRENT APPLICATION NUMBER: US/09/783,590 | | | |
| ; CURRENT FILING DATE: 2000-02-15 | | | |
| ; PRIOR APPLICATION NUMBER: 08/420,856 | | | |
| ; PRIOR FILING DATE: 1995-04-12 | | | |
| ; PRIOR APPLICATION NUMBER: 08/346,731 | | | |
| ; PRIOR FILING DATE: 1994-11-21 | | | |
| ; NUMBER OF SEQ ID NOS: 12485 | | | |
| ; SOFTWARE: PatentIn Ver. 2.0 | | | |
| ; SEQ ID NO 10392 | | | |
| ; LENGTH: 478 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Homo sapiens | | | |
| ; FEATURE: | | | |
| ; NAME/KEY: misc feature | | | |
| ; LOCATION: (135) | | | |
| ; OTHER INFORMATION: n equals a,t,g, or c | | | |
| ; NAME/KEY: misc feature | | | |
| ; LOCATION: (139) | | | |
| ; OTHER INFORMATION: n equals a,t,g, or c | | | |

```

RESULT 11
US-10-029-386-22674/c
; Sequence 2674, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22674
; LENGTH: 644
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004501.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = .2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: NT HIT: AF041260.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: O75363, EVALUE 6.00e-79
; OTHER INFORMATION: EST_HUMAN HIT: B1549901.1, EVALUE 0.00e+00
US-10-029-386-22674

Query Match          19.0%; Score 582.6; DB 16; Length 644;
Best Local Similarity 99.3%; Pred. No. 1.3e-152;
Matches 585; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```
; NAME/KEY: misc feature
; LOCATION: (276)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (281)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (334)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (451)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-10392

Query Match      13.0%; Score 398.6; DB 9; Length 478;
Best Local Similarity 97.0%; Pred. No. 5.1e-101;
Matches 456; Conservative 0; Mismatches 9; Indels 5; Gaps 5;

Qy 141 CCGACAGCCAGGCACTGGGCGACGACGCACTGGAGACCCAGGACCCCTGTGCGAGGAGCAGCT 200
Db 9 CCGACAGCCAGGCACTGGGCGACGCACTGGAGACCCAGGACCCCTGTGCGAGGAGCAGCT 68

Qy 201 CCGGCTGACAGAGGGGACTGAAGATCTCCACAGGGGGCTCAGCAGAGCAATGGGTAA 260
Db 69 CCGGCTGACAGAGGGGACTGAAGATCTCCACAGGGGGCTCAGCAGAGCAATGGGTAA 128

Qy 261 CCAATAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGATGAACCCAGAGCAGAGACTTA 320
Db 129 CCAATNAGTNTTCCCAAGAGTTGAAGACCAAGAGATGAACCCAGAGCAGAGACTTA 188

Qy 321 CCAGGACAAACGGCTGCTCTGAAACGGGGTTCAGTGGTGGTTCGACCCACACAGTTCA 380
Db 199 CCAGGACAAACGGCTGCTCTGAAACGGGGTTCAGTGGTGGTTCGACCCACACAGTTCA 248

Qy 381 GCACCTAGAGAGTGCAGCTTGGGAAT-AAGTGTCAAGACGGATTAATGTGCCACTTCTT 439
Db 249 GCACCTAGAGAGTGCAGCTTGGGAATNAAAGTNTCAAGACGGATTAATGTGCCACTTCTT 308

Qy 440 CCCCCGA-GACAAACGGAG-TAAGTGTGTTGCGGATCCCAACGGAAGAATCTTGGGAA 497
Db 309 CCCCCGAGACACCGAGATTAAAGTNCITGTCGGATGCCAAGGAAGAATCTTGGGAA 368

Qy 498 AGAGGCCAAACCCGAGGACCAAGCTGCTGTAATCTCGTTTTTCTTTGATGCTCTCTCGGCC 557
Db 369 AGAGGCCAAACCCGAGGACCAAGCTGCTGTAATCTCGTTTTTCTTTGATGCTCTCTCGGCC 428

Qy 558 TG-ThACAGGAGTACCGGA-GACCAAGCCGAGATTCAATCCCTTGGATC 605
Db 429 TGTACCAGGAGTACCGGAGNCCAAAGCCGAGGTTCAATTCCTTGGTTC 478

RESULT 13
US-10-029-386-3914/c
; Sequence 3914, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn. Sharron G.
; APPLICANT: Rank, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3914
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.93
```

```
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.75
; OTHER INFORMATION: SWISSPROT HIT: O75363, EVALUE 3.00e-06
; OTHER INFORMATION: NT HIT: AF041260.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE825197.1, EVALUE 0.00e+00
US-10-029-386-3914

Query Match      10.9%; Score 332.8; DB 16; Length 576;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 334; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1929 CTTGGGACCAAAAGCGGATGTTGGATGCTCAAGTCGCAACAGACCCAGTATCCATCGGACC 1988
Db 336 CCAGGGACCAAAAGCGGATGTTGGATGCTCAAGTCGCAACAGACCCAGTATCCATCGGACC 277

Qy 1989 AGTTGGCAAAACCCCAAGTAACAAATCAGCAGGTTCCACACAGGTTCTCTGCCACCAAG 2048
Db 276 AGTTGGCAAAACCCCAAGTAACAAATCAGCAGGTTCCACACAGGTTCTCTGCCACCAAG 217

Qy 2049 ATGTGTTCTCTTACTCCATCTCTCTCCCAAAACAGCTCCATGTATATATATTTCTTGATG 2108
Db 216 ATGTGTTCTCTTACTCCATCTCTCTCCCAAAACAGCTCCATGTATATATATTTCTTGATG 157

Qy 2109 GCCAGCAAAATGAATTTCTGCTAGAAATTAAGCCGAGCTGTTGTATATTGAGGTGTTATT 2168
Db 156 GCCAGCAAAATGAATTTCTGCTAGAAATTAAGCCGAGCTGTTGTATATTGAGGTGTTATT 97

Qy 2169 ATTTACGTCCTGCTGTCAGTCTTTTCTGGCAAAATAACAGATAAGATGTTTAGCAGGTCA 2228
Db 96 ATTTACGTCCTGCTGTCAGTCTTTTCTGGCAAAATAACAGATAAGATGTTTAGCAGGTCA 37

Qy 2229 CCTAGTTGGGTCAAGAGATCGATGATCACCAGCA 2264
Db 36 CCTAGTTGGGTCAAGAGATCGATGATCACCAGCA 1

RESULT 14
US-10-800-322-327/c
; Sequence 327, Application US/10800322
; Publication No. US20050053967A1
; GENERAL INFORMATION:
; APPLICANT: MEDIMOLECULAR PTY LTD
; TITLE OF INVENTION: NOVEL MARKERS AND USES THEREOF
; FILE REFERENCE: 2558321/TDO
; CURRENT APPLICATION NUMBER: US/10/800,322
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US 60/322228
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 327
; LENGTH: 372
; TYPE: DNA
; ORGANISM: mammalian
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: ( )
; OTHER INFORMATION: "n" is an unknown nucleotide
US-10-800-322-327

Query Match      10.7%; Score 328; DB 21; Length 372;
Best Local Similarity 97.6%; Pred. No. 3e-81;
Matches 328; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2625 AGTTATAGACAGAAATGCTCTAGAGGAATTAAGCGTTTCTAGTAACGGTCTCATC 2684
Db 356 AGTTATAGACAGAAATGCTCTAGAGGAATTAAGCGTTTCTAGTAACGGTCTCATC 297

Qy 2685 CTCTAAGCTTCAAGAAAGGAGACGAAATCCATTTGTTTAAATTCACATCTCAAGGAGG 2744
Db 296 CTNTAAGCTTGAAGAAAGGAGACGAAATCCNTTTGTTTAAATTCACATNTCAAGGAGG 237

Qy 2745 AGAACCCGGGCTGTGTGGGTGGTTCACATTTCTCTAGAACGGAATGTGTGGGTATAGA 2804
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:48:07 ; Search time 13508.8 seconds
(without alignments)
11428.001 Million cell updates/sec

Title: US-08-731-499-10
Perfect score: 3186
Sequence: 1 ATGCAATCGAAAGTGACAGG.....GGAACACTACAGTTGTGTAA 3186

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hhg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 3183.6 | 99.9 | 3186 | 6 | BD085734 Genes fro |
| 2 | 3040.6 | 95.4 | 5632 | 6 | AR157086 Sequence |
| 3 | 3040.6 | 95.4 | 5632 | 9 | AF041259 Homo sapi |
| 4 | 3029.6 | 95.1 | 5633 | 6 | CQ726295 Sequence |
| 5 | 1561.2 | 49.0 | 128871 | 9 | AL157838 Human DNA |
| 6 | 1559.6 | 49.0 | 10365 | 6 | BD085733 Genes fro |
| 7 | 1559.6 | 49.0 | 121143 | 9 | AF312915 Homo sapi |
| 8 | 1556.4 | 48.9 | 105023 | 2 | AC116668 Trypanoso |
| 9 | 1106.2 | 34.7 | 1507 | 6 | AR070326 Sequence |
| 10 | 1106.2 | 34.7 | 1507 | 6 | BD085727 Genes fro |
| 11 | 733 | 23.0 | 114285 | 10 | AL844576 Mouse DNA |
| 12 | 722.6 | 22.7 | 208979 | 2 | AC023610 Mus muscu |
| 13 | 716.2 | 22.5 | 184420 | 2 | AC118373 Rattus no |
| 14 | 716.2 | 22.5 | 246332 | 2 | AC099079 Rattus no |
| 15 | 444.4 | 13.9 | 469 | 6 | BD060160 Secreted |
| 16 | 315.8 | 9.9 | 2750 | 5 | AF419155 Xenopus l |
| 17 | 315.6 | 9.9 | 3512 | 5 | BC072191 Xenopus l |
| 18 | 274.4 | 8.6 | 1196 | 10 | BC046393 Mus muscu |
| 19 | 113.6 | 3.6 | 212026 | 2 | AC136666 Rattus no |

| | | | | | | | |
|---|----|-------|-----|--------|----|----------|--------------------|
| c | 20 | 113.6 | 3.6 | 244851 | 2 | AC094607 | AC094607 Rattus no |
| c | 21 | 113.6 | 3.6 | 268559 | 2 | AC112808 | AC112808 Rattus no |
| c | 22 | 104.6 | 3.3 | 4934 | 10 | BC053104 | BC053104 Mus muscu |
| c | 23 | 104.6 | 3.3 | 175110 | 10 | AC142100 | AC142100 Mus muscu |
| c | 24 | 104 | 3.3 | 5982 | 6 | CQ714295 | CQ714295 Sequence |
| c | 25 | 104 | 3.3 | 112797 | 9 | AC018413 | AC018413 Homo sapi |
| c | 26 | 104 | 3.3 | 178717 | 2 | AC024396 | AC024396 Homo sapi |
| c | 27 | 104 | 3.3 | 211981 | 9 | AC009716 | AC009716 Homo sapi |
| c | 28 | 102.4 | 3.2 | 6033 | 9 | D86975 | D86975 Homo sapien |
| c | 29 | 97 | 3.0 | 8511 | 5 | BC081149 | BC081149 Xenopus l |
| c | 30 | 79.8 | 2.5 | 4265 | 10 | BC033594 | BC033594 Mus muscu |
| c | 31 | 79.8 | 2.5 | 4417 | 10 | AK122270 | AK122270 Mus muscu |
| c | 32 | 79.8 | 2.5 | 195922 | 10 | AC140213 | AC140213 Mus muscu |
| c | 33 | 77.8 | 2.4 | 243887 | 2 | AC024898 | AC024898 Homo sapi |
| c | 34 | 76.2 | 2.4 | 165094 | 2 | CR547124 | CR547124 Danio rer |
| c | 35 | 76.2 | 2.4 | 174225 | 5 | BX005071 | BX005071 Zebrafish |
| c | 36 | 76.2 | 2.4 | 192578 | 5 | AL732455 | AL732455 Zebrafish |
| c | 37 | 75.8 | 2.4 | 256281 | 2 | AC094062 | AC094062 Rattus no |
| c | 38 | 75.8 | 2.4 | 266956 | 2 | AC133822 | AC133822 Rattus no |
| c | 39 | 72 | 2.3 | 4935 | 6 | CQ718121 | CQ718121 Sequence |
| c | 40 | 72 | 2.3 | 4935 | 9 | AB002388 | AB002388 Homo sapi |
| c | 41 | 72 | 2.3 | 163120 | 9 | AC011454 | AC011454 Homo sapi |
| c | 42 | 71.2 | 2.2 | 171239 | 2 | AC150136 | AC150136 Gallus ga |
| c | 43 | 71.2 | 2.2 | 190305 | 2 | AC150054 | AC150054 Gallus ga |
| c | 44 | 70 | 2.2 | 70 | 6 | CQ876260 | CQ876260 Sequence |
| c | 45 | 66.8 | 2.1 | 1958 | 6 | CQ730003 | CQ730003 Sequence |

ALIGNMENTS

RESULT 1
BD085734
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BD085734
Genes from the 20q13 amplicon and their uses.
BD085734
BD085734.1 GI:22631344
JP 2001524802-A/10.
synthetic construct
other sequences; artificial sequences.
1 (bases 1 to 3186)
Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and
Romms,J.
Patent: JP 2001524802-A 10 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS
Artificial Sequence
PN
JP 2001524802-A/10
PD
04-DEC-2001
PF
15-JUL-1997 JP 1998506264
PR
15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI
JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI

DAVID KOWBEL,
PI
JOHANNA ROMMENS
PC
C12N15/11, C12Q1/68, A61K48/00
CC
Description of Artificial Sequence: ZABC1 Open Reading Frame FH
Key
Location/Qualifiers
FT
source
1..3186
Location/Qualifiers
1..3186
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

Query Match 99.9%; Score 3183.6; DB 6; Length 3186;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3180; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCATCGAAGTGACAGGAAACATGCACTCAATCCCTCTTAATGTACATGATGGG 60

Db 1 ATGCAATCGAAAGTGACAGGAAACATGCAAACTCAATCCCTCTTAATGTACATGATGGG 60
Qy 61 CCAGAGTGATTTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCTTCTCAATG 120
Db 61 CCAGAGTGATTTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCTTCTCAATG 120
Qy 121 AAAGGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTCAATCCAAATCGAG 180
Db 121 AAAGGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTCAATCCAAATCGAG 180
Qy 181 GGGTATATGCCCCTTGGATTTGATGTTCTGAGCAGACCTTTCAACATTCAGAGAAGCTT 240
Db 181 GGGTATATGCCCCTTGGATTTGATGTTCTGAGCAGACCTTTCAACATTCAGAGAAGCTT 240
Qy 241 AATAAATGCTTTAATGCAACACCGGCTACCTCTGTGAACGAGAGTTCTTCGGGTT 300
Db 241 AATAAATGCTTTAATGCAACACCGGCTACCTCTGTGAACGAGAGTTCTTCGGGTT 300
Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 360
Db 301 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 360
Qy 361 AAGAATTCGAAGGAAAAATGAAATTTAGTGTGAGGTATGTGGGCGACATTTAGAGTGCCT 420
Db 361 AAGAATTCGAAGGAAAAATGAAATTTAGTGTGAGGTATGTGGGCGACATTTAGAGTGCCT 420
Qy 421 TTTGATGTTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGAACATG 480
Db 421 TTTGATGTTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGAACATG 480
Qy 481 TGGCGAAGAGMTTSSRSAGCTTGGTCTTAAATAATCAATGCGGACACATAATGGC 540
Db 481 TGGCGAAGAGMTTCAAGAGACCTTGGTCTTAAATAATCAATGCGGACACATAATGGC 540
Qy 541 AAATTCGGGGGCGAGAACAACTGCGAACAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 600
Db 541 AAATTCGGGGGCGAGAACAACTGCGAACAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 600
Qy 601 GAGTCTGTCAGGTGTCAGCGGCGGAGAGCATCTCTCTTCAAAATCTGCATGGTT 660
Db 601 GAGTCTGTCAGGTGTCAGCGGCGGAGAGCATCTCTCTTCAAAATCTGCATGGTT 660
Qy 661 TGTGGCTCTCTATTTCCAAATAAAGAAAGTCTAATTGAGACCGCAAGGTGCACACAAA 720
Db 661 TGTGGCTCTCTATTTCCAAATAAAGAAAGTCTAATTGAGACCGCAAGGTGCACACAAA 720
Qy 721 AAAACTGCTTTCCGTPACAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCC 780
Db 721 AAAACTGCTTTCCGTPACAGCAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCC 780
Qy 781 TCGAGGGAGGACTTCCTGCAGTTGTTCACTTGAACCAAAATCTCAACCTGMAACGGG 840
Db 781 TCGAGGGAGGACTTCCTGCAGTTGTTCACTTGAACCAAAATCTCAACCTGMAACGGG 840
Qy 841 AAGAAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTTCAACCTTCCAGGCTTGGCAG 900
Db 841 AAGAAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTTCAACCTTCCAGGCTTGGCAG 900
Qy 901 CTGGCTACCAAGGAAAAAGTTGCCATTTTCCAAAGAAAGTGAAGGAATCGGGGCAAGAAGG 960
Db 901 CTGGCTACCAAGGAAAAAGTTGCCATTTTCCAAAGAAAGTGAAGGAATCGGGGCAAGAAGG 960
Qy 961 AGCACCAGAACGAGATTCGAGTTCGAGAGGAGCTTGGAGAAACAAATAAGGGGAGT 1020
Db 961 AGCACCAGAACGAGATTCGAGTTCGAGAGGAGCTTGGAGAAACAAATAAGGGGAGT 1020
Qy 1021 TGTGAGGCTCTCCCAAGAGAAAGAGAAAGTGCACCACTCCACGGCGAAGCGCCCTCC 1080
Db 1021 TGTGAGGCTCTCCCAAGAGAAAGAGAAAGTGCACCACTCCACGGCGAAGCGCCCTCC 1080
Qy 1081 GTGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAGGCCACTCACTGCTCCGAGTGC 1140

Db 1081 GTGACGCGGATCCCAAGTTTACCCAGTAGCAAGAGAGGCCACTCACTGCTCCGAGTGC 1140
Qy 1141 GGCAAAAGCTTTCAAGAACTACACACAGCTGTCTTGCACCTCCAGGGTCCACAAGAGGAC 1200
Db 1141 GGCAAAAGCTTTCAAGAACTACACACAGCTGTCTTGCACCTCCAGGGTCCACAAGAGGAC 1200
Qy 1201 CGGAGGCGCGCGCGAGTCCGCCACCATGTTCTGTGAACGAGGAGCGCGGGAAGTGT 1260
Db 1201 CGGAGGCGCGCGCGAGTCCGCCACCATGTTCTGTGAACGAGGAGCGCGGGAAGTGT 1260
Qy 1261 TCTCTGACCTCGCGGCGCTCTGATGAAATCGAGCGCTGGATCGAGGGGAAGTGT 1320
Db 1261 TCTCTGACCTCGCGGCGCTCTGATGAAATCGAGCGCTGGATCGAGGGGAAGTGT 1320
Qy 1321 TCTGAAGACGATCTGAGGATGGCTTCCGAGGAATCCATCTGGATAAAAATGATCAT 1380
Db 1321 TCTGAAGACGATCTGAGGATGGCTTCCGAGGAATCCATCTGGATAAAAATGATCAT 1380
Qy 1381 GGAGGAAAAATAAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGTGGAAGTTTTTC 1440
Db 1381 GGAGGAAAAATAAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGTGGAAGTTTTTC 1440
Qy 1441 CGTTCAAATTTATTAACCTCAATATTTCAATCTCAGAACGATACAGGTGAAAAACCATACAAA 1500
Db 1441 CGTTCAAATTTATTAACCTCAATATTTCAATCTCAGAACGATACAGGTGAAAAACCATACAAA 1500
Qy 1501 TGTGAATTTTGTGAATATGCTGACGCCAGAGACATCTCTGAGGTATCACTTGGAGAGA 1560
Db 1501 TGTGAATTTTGTGAATATGCTGACGCCAGAGACATCTCTGAGGTATCACTTGGAGAGA 1560
Qy 1561 CATCAAGGAAAAACAAACCGATGTTGCTGCTGAAAGTCAAGAACGATGGTAAAAAATCAG 1620
Db 1561 CATCAAGGAAAAACAAACCGATGTTGCTGCTGAAAGTCAAGAACGATGGTAAAAAATCAG 1620
Qy 1621 GACACTGAAGATGCACTATTAACCGCTGACAGTGCAGAACCAAAATTTGAAAGATTT 1680
Db 1621 GACACTGAAGATGCACTATTAACCGCTGACAGTGCAGAACCAAAATTTGAAAGATTT 1680
Qy 1681 TTTGATGTCGCAAGATGTTTACAGGAGTCCACTGCAAGAGAGCTTAAAGAGATGCT 1740
Db 1681 TTTGATGTCGCAAGATGTTTACAGGAGTCCACTGCAAGAGAGCTTAAAGAGATGCT 1740
Qy 1741 TCTGTTTTTTCAGAAATGTTCTGGGCGAGCTGCTCTCTCACAGCACACAAAGATACTCAG 1800
Db 1741 TCTGTTTTTTCAGAAATGTTCTGGGCGAGCTGCTCTCTCACAGCACACAAAGATACTCAG 1800
Qy 1801 GATTTCCATAAAATGACAGTGTGATGAAAGTGAATAAAACCCCTACCCCT 1860
Db 1801 GATTTCCATAAAATGACAGTGTGATGAAAGTGAATAAAACCCCTACCCCT 1860
Qy 1861 GCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAACCTCATC 1920
Db 1861 GCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAACCTCATC 1920
Qy 1921 TGTAGAACAAGGCGGATGTTTACTCTCTCCGATGGCAGTACACCCATAACCTTGAA 1980
Db 1921 TGTAGAACAAGGCGGATGTTTACTCTCTCCGATGGCAGTACACCCATAACCTTGAA 1980
Qy 1981 GTTAGCCCCCAAGAGAAAGCAACCGAGACCGAGCTGACTGATGAAAGTGAATAAAACCCCTACCCCT 2040
Db 1981 GTTAGCCCCCAAGAGAAAGCAACCGAGACCGAGCTGACTGATGAAAGTGAATAAAACCCCTACCCCT 2040
Qy 2041 GATTTGTCAGAAAAACCTTTTAAATTTATTCGTTGGGGCTCTTCAAAATGGCCCGCAATT 2100
Db 2041 GATTTGTCAGAAAAACCTTTTAAATTTATTCGTTGGGGCTCTTCAAAATGGCCCGCAATT 2100
Qy 2101 TCTTTGAGTAAAAAGTTTGATTTCAAGTATACCTGTCCATTTTGTGACTTCAAGCAATTT 2160
Db 2101 TCTTTGAGTAAAAAGTTTGATTTCAAGTATACCTGTCCATTTTGTGACTTCAAGCAATTT 2160
Qy 2161 TATCCAGAGTTTAAATGATGACACAGAGCTGGAGCATTAATAAATCTCTGACGTTTCAAT 2220
Db 2161 TATCCAGAGTTTAAATGATGACACAGAGCTGGAGCATTAATAAATCTCTGACGTTTCAAT 2220

```
QY 2221 AAAAAGTCTGAAACAGTCTCTGCTTAGAGTCGACGTACCGATGCCCGCAGCGTTG 2280
DB 2221 AAAAAGTCTGAAACAGTCTCTGCTTAGAGTCGACGTACCGATGCCCGCAGCGTTG 2280
QY 2281 CTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTC 2340
DB 2281 CTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTC 2340
QY 2341 CCGGGCAGTCCAAATCCCTGCTCATCTCGAAGGGGAAAGCAGAGCCCTCTCGGGCAGGC 2400
DB 2341 CCGGGCAGTCCAAATCCCTGCTCATCTCGAAGGGGAAAGCAGAGCCCTCTCGGGCAGGC 2400
QY 2401 AAGGCCCTCTGACTTCAGGATGACCTAGCCTTTAGCCCAAGTAACCTGAGTCC 2460
DB 2401 AAGGCCCTCTGACTTCAGGATGACCTAGCCTTTAGCCCAAGTAACCTGAGTCC 2460
QY 2461 CACAGACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACAGGCAACAGCAATCTGAG 2520
DB 2461 CACAGACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACAGGCAACAGCAATCTGAG 2520
QY 2521 ATGTTTCTAAACCAAGTGTTCCTCTGCACCGGATAAGACAAAGACCCGAGACAAA 2580
DB 2521 ATGTTTCTAAACCAAGTGTTCCTCTGCACCGGATAAGACAAAGACCCGAGACAAA 2580
QY 2581 TTGAAACCTCTTCAGTAGCTCTCTCAGGCCACCTCTGGCAGCAGTAACATCATGGT 2640
DB 2581 TTGAAACCTCTTCAGTAGCTCTCTCAGGCCACCTCTGGCAGCAGTAACATCATGGT 2640
QY 2641 TCCATCGACTACCCCGCAAGAACGACAGCCGTCGACCTCCGGGAAGAGACTATTTC 2700
DB 2641 TCCATCGACTACCCCGCAAGAACGACAGCCGTCGACCTCCGGGAAGAGACTATTTC 2700
QY 2701 TGTAAATCGAGTCCAGCAATCTGCAGCAAGATTTGTGAGCCCTTCCAAAAGACTG 2760
DB 2701 TGTAAATCGAGTCCAGCAATCTGCAGCAAGATTTGTGAGCCCTTCCAAAAGACTG 2760
QY 2761 AAGTCCAGCTGGTGGCTTGAAGTTCAGAGGCAATCAGAGCCCGGCAATTCAGAGGCTAT 2820
DB 2761 AAGTCCAGCTGGTGGCTTGAAGTTCAGAGGCAATCAGAGCCCGGCAATTCAGAGGCTAT 2820
QY 2821 GACCTTCCCAAGTACCATATGTCAGAGGCAATCAGAGGCTATCAGAGGCTAT 2880
DB 2821 GACCTTCCCAAGTACCATATGTCAGAGGCAATCAGAGGCTATCAGAGGCTAT 2880
QY 2881 TATCCGTCCGAGGCGTCCCTCCCAACCAAGTTCTCTGAGCTCCAGCGAGTCTGATTCT 2940
DB 2881 TATCCGTCCGAGGCGTCCCTCCCAACCAAGTTCTCTGAGCTCCAGCGAGTCTGATTCT 2940
QY 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGCTCCGGGCCACTTTACATTTGTTG 3000
DB 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGCTCCGGGCCACTTTACATTTGTTG 3000
QY 3001 CCTGCTGTGCTCAGAGTCCAGCTCAGCTTAGAGGCTTTGTTGAGTTCAGTCTTA 3060
DB 3001 CCTGCTGTGCTCAGAGTCCAGCTCAGCTTAGAGGCTTTGTTGAGTTCAGTCTTA 3060
QY 3061 CTCCCAAGAAATTAATTTTACTTCTATCTTTCAGAGGCAATGGTGAAGTCTACTGAA 3120
DB 3061 CTCCCAAGAAATTAATTTTACTTCTATCTTTCAGAGGCAATGGTGAAGTCTACTGAA 3120
QY 3121 ATAAGCTGTGATGTGATCATATAAATCATATAGGAATCTGCAAGGAAACATACAGTT 3180
DB 3121 ATAAGCTGTGATGTGATCATATAAATCATATAGGAATCTGCAAGGAAACATACAGTT 3180
QY 3181 GTGTAA 3186
DB 3181 GTGTAA 3186
```

RESULT 2
AR157086
LOCUS

```
DEFINITION Sequence 3 from patent US 6242590.
ACCESSION AR157086
VERSION AR157086.1 GI:15125790
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5632)
AUTHORS Cowbert, L.M.
TITLE Antisense modulation of zinc finger protein-217 expression
JOURNAL Patent: US 6242590-A 3 05-JUN-2001;
FEATURES
    Location/Qualifiers
        source
            1..5632
            /organism="unknown"
            /mol_type="unassigned DNA"
ORIGIN
    Query Match          95.4%; Score 3040.6; DB 6; Length 5632;
    Best Local Similarity 95.8%; Pred. No. 0;
    Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;
QY 1 ATGCAATCGAAAGTGCACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGG 60
DB 272 ATGCAATCGAAAGTGCACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGG 331
QY 61 CCAGAAGTGAATGGAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 120
DB 332 CCAGAAGTGAATGGAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 391
QY 121 AAAGGGACCGCTGTTTCCATTCGGAGTACACAAAGAAANAATGTCCATCCAAATCGAG 180
DB 392 AAAGGGACCGCTGTTTCCATTCGGAGTACACAAAGAAANAATGTCCATCCAAATCGAG 451
QY 181 GGGTATATGCCCTTGGATTTGCATGTTCTGCAGCCAGACCTTTCACACATTCAGAAGACCTT 240
DB 452 GGGTATATGCCCTTGGATTTGCATGTTCTGCAGCCAGACCTTTCACACATTCAGAAGACCTT 511
QY 241 AATAAATGTCCTTAATGCAACACCGGCTACCTCTGTGAACCGAGAGTCTTTCGGGTT 300
DB 512 AATAAATGTCCTTAATGCAACACCGGCTACCTCTGTGAACCGAGAGTCTTTCGGGTT 571
QY 301 GAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTCCGAAACAGAACCTCCCAAGGAA 360
DB 572 GAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTCCGAAACAGAACCTCCCAAGGAA 631
QY 361 AAGAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGAGACATTTAGAGTCGT 420
DB 632 AAGAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGAGACATTTAGAGTCGT 691
QY 421 TTTGATGTTGAGATCCCATGAGAACACACAAAGATTTCTTCACTTACGCGGTGTAACATG 480
DB 692 TTTGATGTTGAGATCCCATGAGAACACACAAAGATTTCTTCACTTACGCGGTGTAACATG 751
QY 481 TGCGGAAGAGGTTTSSSSAGCCCTTGTCTTAAATACATGCGGACACATTAATGGC 540
DB 752 TGCGGAAGAGGTTTSSSSAGCCCTTGTCTTAAATACATGCGGACACATTAATGGC 811
QY 541 AAATTCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 600
DB 812 AAATTCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 871
QY 601 GAGTCTGTCCAGGTGCAAGCGCCGAGAGCATCTCTCTCTTACAAAATCTGCAATGTT 660
DB 872 GAGTCTGTCCAGGTGCAAGCGCCGAGAGCATCTCTCTCTTACAAAATCTGCAATGTT 931
QY 661 TGTGGCTCTCTTATTTCCAAATAAAGAGTCTTAATTTAGACCCGCAAGTGCACACCAA 720
DB 932 TGTGGCTCTCTTATTTCCAAATAAAGAGTCTTAATTTAGACCCGCAAGTGCACACCAA 991
QY 721 AAAAATGCTTTCCGTTACCCAGCAGCGCAGACAGACTCTCCACAGGAGGAAATGCCGTCC 780
DB 992 AAAAATGCTTTCCGTTACCCAGCAGCGCAGACAGACTCTCCACAGGAGGAAATGCCGTCC 1051
```

QY 781 TCAGAGGAGACTTCTCAGTGTGTTCAACTTGAGACCAAAATCTCACTGAAACGGGG 840
DB TCGAGGGAGAGACTTCTCAGTGTGTTCAACTTGAGACCAAAATCTCACTTGAACGGGG 1111
QY 841 AAGAAGCCTGTCCAGATGATCCCTCAGCTCGATCCGTTTCAACAAGCTTCCAGGCTTGGCAG 900
DB AAGAAGCCTGTCCAGATGATCCCTCAGCTCGATCCGTTTCAACAAGCTTCCAGGCTTGGCAG 1171
QY 901 CTGGCTACCAAGAGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAGGG 960
DB CTGGCTACCAAGAGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAGGG 1231
QY 961 AGCACCGACACGAGCTTCCAGTTCGAGAGAGAGCTTGGAGAAACAAATAAGGGCAGT 1020
DB AGCACCGACACGAGCTTCCAGTTCGAGAGAGAGCTTGGAGAAACAAATAAGGGCAGT 1291
QY 1021 TGTGAGGCCCTTCGCAAGAGAAAGAGAGTGCAAAACATCCCGAGGGGAAAGCGCCCTCC 1080
DB TGTGAGGCCCTTCGCAAGAGAAAGAGAGTGCAAAACATCCCGAGGGGAAAGCGCCCTCC 1351
QY 1081 GTGGACGGGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGC 1140
DB GTGGACGGGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGC 1411
QY 1141 GGCAAAGCTTTCAGAACCTACACAGCTGGTCTTGGCACTCCAGGGTCCACAAGAGGAC 1200
DB GGCAAAGCTTTCAGAACCTACCAACAGCTGGTCTTGGCACTCCAGGGTCCACAAGAGGAC 1471
QY 1201 CGGAGGGCGGCGGAGTGCGCCACATGCTGTGGACGGGAGCGACCGGGGACGTTG 1260
DB CGGAGGGCGGCGGAGTGCGCCACATGCTGTGGACGGGAGCGACCGGGGACGTTG 1531
QY 1261 TCTCTGACCTCGCGCCCTCTGGATGAAATCGAGCCGTGGATCGAGGGGAGGTTG 1320
DB TCTCTGACCTCGCGCCCTCTGGATGAAATCGAGCCGTGGATCGAGGGGAGGTTG 1591
QY 1321 TCTGAAGACGGATCTGAGGATGGGCTTCCGAGGAATCCATCTGGATAAAATATGATGAT 1380
DB TCTGAAGACGGATCTGAGGATGGGCTTCCGAGGAATCCATCTGGATAAAATATGATGAT 1651
QY 1381 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATGTGGAAAGTTTTTC 1440
DB GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATGTGGAAAGTTTTTC 1711
QY 1441 CGTTCAAAATTTACCTCAATTTTCACTCTCAAGCGCATACAGGTGAAAACCCATACAAA 1500
DB CGTTCAAAATTTACCTCAATTTTCACTCTCAAGCGCATACAGGTGAAAACCCATACAAA 1771
QY 1501 TGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCTGAGGTATCACTTGGAGAGA 1560
DB TGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCTGAGGTATCACTTGGAGAGA 1831
QY 1561 CATCAAGGAAAAACAAACCGATGTTGCTGTGAAGTCAAGAACGATGGTAAAAATCAG 1620
DB CATCAAGGAAAAACAAACCGATGTTGCTGTGAAGTCAAGAACGATGGTAAAAATCAG 1891
QY 1621 GACACTGAAGTGCATTAACCCCTGCAGTGCAGAAACCAAAATTTTGAARAAGTTT 1680
DB GACACTGAAGTGCATTAACCCCTGCAGTGCAGAAACCAAAATTTTGAARAAGTTT 1951
QY 1681 TTTGATGTTGCAAAAGATGTTACAGGCAGTCCACCTGCAAAAGCAGCTTAAAGGAGATGCT 1740
DB TTTGATGTTGCAAAAGATGTTACAGGCAGTCCACCTGCAAAAGCAGCTTAAAGGAGATGCT 2011
QY 1741 TCTGTTTTTCAAAATGTTTGGGACAGGCTGTCTCTCAACAGCAACAAAGATACTCAG 1800
DB TCTGTTTTTCAAAATGTTTGGGACAGGCTGTCTCTCAACAGCAACAAAGATACTCAG 2071
QY 1801 GATTTCCATTAATAATGCAGCTGATCAAGTGTCTGATTAAGTAATAAAACCCCTACCCCT 1860
DB GATTTCCATTAATAATGCAGCTGATCAAGTGTCTGATTAAGTAATAAAACCCCTACCCCT 2131
QY 1861 GCTTACCTGGACCTGTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAACCTCATC 1920

DB 2132 GCTTACCTGGACCTGTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAACCTCATC 2191
QY 1921 TGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAGTACACCCATAACCTTGAA 1980
DB TGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAGTACACCCATAACCTTGAA 2251
QY 1981 GTTAGCCCCAAAGAGAACCAACGAGACCGCAGCTGACTGCAGATACAGGCCCAAGTGTG 2040
DB GTTAGCCCCAAAGAGAACCAACGAGACCGCAGCTGACTGCAGATACAGGCCCAAGTGTG 2311
QY 2041 GATTGTCAAGAAAACTTTTAAATTTATCCGTGGGGCTCTTCAATTTGCCGGCAATTT 2100
DB GATTGTCAAGAAAACTTTTAAATTTATCCGTGGGGCTCTTCAATTTGCCGGCAATTT 2371
QY 2101 TCTTTGAGTAAAGCTTTGATTCGAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTT 2160
DB TCTTTGAGTAAAGCTTTGATTCGAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTT 2431
QY 2161 TATCCAGAGTTTAAATGATGACACAGAGACTGGAGCATAAATACAACTCTGACGTTTCA 2220
DB TATCCAGAGTTTAAATGATGACACAGAGACTGGAGCATAAATACAACTCTGACGTTTCA 2491
QY 2221 AAAAACTGTGAAACAAAGTCTCTTGTAGAAAGTCAAGTACCGGATGCCCGCAGCGTTG 2280
DB AAAAACTGTGAAACAAAGTCTCTTGTAGAAAGTCAAGTACCGGATGCCCGCAGCGTTG 2551
QY 2281 CTGGAAAAAGATGTCCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2340
DB CTGGAAAAAGATGTCCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2611
QY 2341 CCGCGCAGTCCAAATCCCTGCCATCTCGAAGGGAGCAGAGCCCTCTCTGGGCGAGGC 2400
DB CCGCGCAGTCCAAATCCCTGCCATCTCGAAGGGAGCAGAGCCCTCTCTGGGCGAGGC 2671
QY 2401 AAGGCCCTCTGACTTCAAGGATAGACTCTAGCACTTTAGCCCCCAAGTAACTGAAGTCC 2460
DB AAGGCCCTCTGACTTCAAGGATAGACTCTAGCACTTTAGCCCCCAAGTAACTGAAGTCC 2731
QY 2461 CACAGACACAGCAGATGTTGGGGTCCAGGGGCGCCACGAGCAACCAATCTGAG 2520
DB CACAGACACAGCAGATGTTGGGGTCCAGGGGCGCCACGAGCAACCAATCTGAG 2791
QY 2521 ATGTTTCTTAAACCAAGTGTTCCTGCAACGGATAGACAAAAGAACCCGAGACAAA 2580
DB ATGTTTCTTAAACCAAGTGTTCCTGCAACGGATAGACAAAAGAACCCGAGACAAA 2851
QY 2581 TTGAAACCTTTCAGTAGTCTCTTCTCAGCCACCTCTCGGAGCAGTAAATCAATGTT 2640
DB TTGAAACCTTTCAGTAGTCTCTTCTCAGCCACCTCTCGGAGCAGTAAATCAATGTT 2911
QY 2641 TCCATCGACTACCCCGCAAGAACGACAGCCCTGGGCACTCTCGGGAGAGACTATTTC 2700
DB TCCATCGACTACCCCGCAAGAACGACAGCCCTGGGCACTCTCGGGAGAGACTATTTC 2971
QY 2701 TGTAACTCGGAGTGCAGCAATACTGACAGAAATTTGTGAGCCCTTCCAAAAAGACTG 2760
DB TGTAACTCGGAGTGCAGCAATACTGACAGAAATTTGTGAGCCCTTCCAAAAAGACTG 3031
QY 2761 AAGTCCAGCTGTTGCCCTTGAAGTTCAGCCCGGGGCAATTAAGAGAGGCTAT 2820
DB AAGTCCAGCTGTTGCCCTTGAAGTTCAGCCCGGGGCAATTAAGAGAGGCTAT 3091
QY 2821 GACCTTCCAGTACCATATGTCAGAGCATCACATCACTGTTTACCCAGGACTGTG 2880
DB GACCTTCCAGTACCATATGTCAGAGCATCACATCACTGTTTACCCAGGACTGTG 3151
QY 2881 TATCCGTGCGAGGCGCTCCCTCCCAACCAAGTTTCTGAGCTCCAGGAGGCTCGATTCT 2940
DB TATCCGTGCGAGGCGCTCCCTCCCAACCAAGTTTCTGAGCTCCAGGAGGCTCGATTCT 3211
QY 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGCTCCGGGCACTTTTACCTTGTG 3000

Db 3212 CCAAAATGCTGACTGTTTCAGAAAGCCCTATGCTGGCTCGGCCACTTTTACACTTGTGTG 3271
Qy 3001 CCGTCTGGTAGTCCAGCATCCAGCTCGAGTTAG----- 3034
Db 3272 CCTGCTGGTAGTCCAGCATCCAGCTCGAGTTAGAAAGGAAAAAGGCTGTGTGCATATCAA 3331
Qy 3035 ----- 3034
Db 3332 CACTTATCTAACAGCATGCCACAAAGAGAAACTATGAGATTTTATGGGAATGCACAT 3391
Qy 3035 -----AAGGCTTGTGGTGG 3047
Db 3392 TATGACCAAAATGACAAAAAACTTGATTCACATAATAGGGGAAAAAGGTTCTTGGTGG 3451
Qy 3048 ATGTCAGTCTTACTCCCATCAATTAATTTTACTTTCATCTTGTGAGAACGAATGGT 3107
Db 3452 ATGTGAGTCTTACTCCCATCAATTAATTTTACTTTCATCTTGTGAGAACGAATGGT 3511
Qy 3108 GAAAGCTACTGAAATAGCTGTGATTTGACTGTACATAAAACATATGAGGAATCTGCAAG 3167
Db 3512 GAAAGCTACTGAAATAGCTGTGATTTGACTGTACATAAAACATATGAGGAATCTGCAAG 3571
Qy 3168 GAACTACTAGTTGTGTAA 3186
Db 3572 GAACTACTAGTTGTGTAA 3590

RESULT 3
AF041259
LOCUS
DEFINITION Homo sapiens breast cancer putative transcription factor (ZABC1)
ACCESSION AF041259
VERSION AF041259.1 GI:3335396
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 5632)
Collins,C., Rommens,J.M., Kowbel,D., Godfrey,T., Tanner,M.,
Hwang,S.-I., Polikoff,D., Nonet,G., Cochran,J., Myambo,K.,
Jay,K.E., Froula,J., Cloutier,T., Kuo,W.-L., Yaswen,P., Dairkee,S.,
Giovanela,J., Hutchinson,G.B., Isola,J., Kallioniemi,O.-P.,
Palazzo,M., Martin,C., Ericsson,C., Pinkel,D., Albertson,D.,
Li,W.-B. and Gray,J.W.
Positional cloning of ZNF217 and NABCl: genes amplified at 20q13.2
and overexpressed in breast carcinoma
Proc. Natl. Acad. Sci. U.S.A. 95 (15), 8703-8708 (1998)
98337979
9671742
2 (bases 1 to 5632)
Collins,C., Rommens,J.M., Kowbel,D., Godfrey,T., Tanner,M.,
Hwang,S.-I., Polikoff,D., Nonet,G., Cochran,J., Myambo,K.,
Jay,K.E., Froula,J., Cloutier,T., Kuo,W.-L., Yaswen,P., Dairkee,S.,
Giovanela,J., Hutchinson,G.B., Isola,J., Kallioniemi,O.-P.,
Palazzo,M., Martin,C., Ericsson,C., Pinkel,D., Albertson,D.,
Li,W.-B. and Gray,J.W.
Direct Submission
Submitted (07-JAN-1998) Life Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
FEATURES
source
1..5632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.2-q13.3"
1..5632
/gene="ZABC1"
272..3418
/gene="ZABC1"
/note="Kruppel-like"

/codon_start=1
/product="breast cancer putative transcription factor"
/protein_id="AAC39895.1"
/db_xref="GI:3335397"
/translation="MQSKVTGNMPTQSLIYMDGPEVIGSSLGSPMEMEDALSMKGTGA
VYPLRATOKNVIQIEGYMPLDMECSOTFTHSEDINKHVLMOHRTLCCEPAVLVVEA
EVLSPLDKSOVQRTPEPKKCNKCNKNEFSCVCGOTFVADPVEIHMRTKDSPTYGCM
CGRRPEWFLKNHMTNTHNGSGARSKLQOGLSSPATINEVVQVHAASISPPYKIC
MYGFLFPNKESLIEHRKVTHTKATFTSSAQTDSFGQMPSSREDFLQLFNLRPKSH
PETGKVPVCIPLDPTFTTQAWQLATKGAICQEVKESQGGSTDDNDSSSEKELG
ETNKGSCAGLSQEKCKSHGEAPSDADPKLPSKCKPCKPCKPCKPCKPCKPCKPCK
HSRVHKDRRGAESPMTSDGROQPTCSPLAAPLDENGAVDRGEGGSEDSGLP
EGIHLDKNDGDKIKHLTSSRECSYCKGFRSNYVNIHLRTHTEKPKYKCFEYAA
AQTSRLRHLERHHEKQTDVAEAVKNGKQDTEALLTADSQTKNLKRFEDGAKD
VTGSPKQKMEPSVFQNLGSAVLSPAHQDQDFHKNAAADDADKNNKNTPTPAYLD
LLKRSASVETANNLICRTKADVTTPPDGTHNLEVPKKEKOTTAADCYRPSVDC
HEKILNLSVGAHNCPAISLSKSLIPSTICPFCTKFTFYEVLMHQRLEHKNYNDVH
KNCRNLSLRRTTCCPALLGKDVPLPSFCCKPKPKPAFPAOSKLSKPAKQSPG
PKAPLTSIGIDSTLAPSNLKSHPQONVGVGAATROOSEMFPKTSVSPADPKTKR
PETKLPPLVAPSPQTLGSSNINGSIDYPAKNDSPWAPGRDYFCNRSASNTAAEFGE
PLPKLKSSVALVDVDPGANYRRGYDLPKYHMRGITSLLPQDCVYPSQALPPKPRF
LSSEVDSVNLTVQKPYGGSGPLYTCVPAGSPASSSTLEGRKPSYQHLNSMAQKR
NYENFIQNAHYRPNDDKT"
470..535
/gene="ZABC1"
/note="zinc finger 1"
559..721
/gene="ZABC1"
/note="zinc finger 2"
744..805
/gene="ZABC1"
/note="zinc finger 3"
923..985
/gene="ZABC1"
/note="zinc finger 4"
1400..1463
/gene="ZABC1"
/note="zinc finger 5"
1688..1750
/gene="ZABC1"
/note="zinc finger 6"
1772..1834
/gene="ZABC1"
/note="zinc finger 7"
2405..2470
/gene="ZABC1"
/note="zinc finger 8"
ORIGIN
Query Match 95.4%; Score 3040.6; DB 9; Length 5632;
Best Local Similarity 95.8%; Pred. NO. 0;
Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;
Qy 1 ATGCAATCGAAGTACAGGAAACATGCCAACTCAATCCCTTAAATGTACATGGATGGG 60
Db 272 ATGCAATCGAAGTACAGGAAACATGCCAACTCAATCCCTTAAATGTACATGGATGGG 331
Qy 61 CCAGAGTGTATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG 120
Db 332 CCAGAGTGTATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG 391
Qy 121 AAAGGACCGCTGTGTGTTTCCATTCGAGCTACACAAGAAAAAATGTATCAATCAATCGAG 180
Db 392 AAAGGACCGCTGTGTGTTTCCATTCGAGCTACACAAGAAAAAATGTATCAATCAATCGAG 451
Qy 181 GGGTATATGCCCTTGGATTGCTGTCGAGCCAGACCTTCACACATTACAGAACCTT 240
Db 452 GGGTATATGCCCTTGGATTGCTGTCGAGCCAGACCTTCACACATTACAGAACCTT 511
Qy 241 AATAAACATGTCTTAATGCAACACCGGCTTACCTCTGTGAAACAGCAGTCTTCCGGTT 300
Db 512 AATAAACATGTCTTAATGCAACACCGGCTTACCTCTGTGAAACAGCAGTCTTCCGGTT 571

| | | | |
|----|------|---|------|
| Qy | 301 | GAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAGTGCAGAACCTCCCAAGGAA | 360 |
| Db | 572 | | |
| | | GAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAGTGCAGAACCTCCCAAGGAA | 631 |
| Qy | 361 | AAGAAATTGCAAGGAAATGAATTTAGCTGTGAGTATGTGGCAGACATTTAGAGTCGCT | 420 |
| Db | 632 | | |
| | | AAGAAATTGCAAGGAAATGAATTTAGCTGTGAGTATGTGGCAGACATTTAGAGTCGCT | 691 |
| Qy | 421 | TTTGATGTTGAGATCCACATGAGAAACACAAAGATTTCTTTCACTTACGGGTGTAACATG | 480 |
| Db | 692 | | |
| | | TTTGATGTTGAGATCCACATGAGAAACACAAAGATTTCTTTCACTTACGGGTGTAACATG | 751 |
| Qy | 481 | TGCGGAAGAAAGMTTSRRSAGCCTTGTTCTTTAAAAATCACATGCGGAACATAATGGC | 540 |
| Db | 752 | | |
| | | TGCGGAAGAAAGATTCAGAGGACCTTGGTTCTTTAAAAATCACATGCGGAACATAATGGC | 811 |
| Qy | 541 | AAATCGGGGGCCAGAACCAACTGCACAGGCTTGGAGAGTAGTCAGCAACGATCAAC | 600 |
| Db | 812 | | |
| | | AAATCGGGGGCCAGAACCAACTGCACAGGCTTGGAGAGTAGTCAGCAACGATCAAC | 871 |
| Qy | 601 | GAGTCCGTCCAGGTGACCGGGCCGAGAGCATCTCCTCTCTTACAAAACTCTGCATGGTT | 660 |
| Db | 872 | | |
| | | GAGTCCGTCCAGGTGACCGGGCCGAGAGCATCTCCTCTCTTACAAAACTCTGCATGGTT | 931 |
| Qy | 661 | TGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA | 720 |
| Db | 932 | | |
| | | TGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA | 991 |
| Qy | 721 | AAAACCTGTTTCGTATCAGCAGCGCGCAGACAGACTCTCCAAGGAGAAATGCCGTCC | 780 |
| Db | 992 | | |
| | | AAAACCTGTTTCGTATCAGCAGCGCGCAGACAGACTCTCCAAGGAGAAATGCCGTCC | 1051 |
| Qy | 781 | TCGAGGGAGGACTTCTCTGCAGTTGTTCAAATTGAGACCAAAATCTCACCTGAAAACGGGG | 840 |
| Db | 1052 | | |
| | | TCGAGGGAGGACTTCTCTGCAGTTGTTCAAATTGAGACCAAAATCTCACCTGAAAACGGGG | 1111 |
| Qy | 841 | AAGAAGCCTGTCCAGATGCATCCCTCAGCTCGATCCGTTCCACCTTCCAGGCTTGGCAG | 900 |
| Db | 1112 | | |
| | | AAGAAGCCTGTCCAGATGCATCCCTCAGCTCGATCCGTTCCACCTTCCAGGCTTGGCAG | 1171 |
| Qy | 901 | CTGCTACCAAGAAAAAGTTGCCATTTGCCAAGAAGTGAAGNAATCGGGGCAAGAGGG | 960 |
| Db | 1172 | | |
| | | CTGCTACCAAGAAAAAGTTGCCATTTGCCAAGAAGTGAAGNAATCGGGGCAAGAGAGGG | 1231 |
| Qy | 961 | AGCACCGAACACGACGATTCCAGTTCCGAGAAGGAGCTTGAGAAAACAAAATAAGGGCAGT | 1020 |
| Db | 1232 | | |
| | | AGCACCGAACACGACGATTCCAGTTCCGAGAAGGAGCTTGAGAAAACAAAATAAGGGCAGT | 1291 |
| Qy | 1021 | TGTGAGGCGCTCTCCGAAGAAAAAGAGATGCAAAACATCCACGCGGAAGCGCCCTCC | 1080 |
| Db | 1292 | | |
| | | TGTGAGGCGCTCTCCGAAGAAAAAGAGATGCAAAACATCCACGCGGAAGCGCCCTCC | 1351 |
| Qy | 1081 | GTGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAAAGCCCACTCACTGCTCCGAGTGC | 1140 |
| Db | 1352 | | |
| | | GTGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAAAGCCCACTCACTGCTCCGAGTGC | 1411 |
| Qy | 1141 | GGCAAGCTTTTCAGAACTCAACACAGCTGTTTGCATCCAGGTTCCAAGAGGAC | 1200 |
| Db | 1412 | | |
| | | GGCAAGCTTTTCAGAACTCAACACAGCTGTTTGCATCCAGGTTCCAAGAGGAC | 1471 |
| Qy | 1201 | CGGAGGGCGGCGGGAGTCGCCACCAATGTTCTGTGACGAGGAGCAGCGGGGACGTGT | 1260 |
| Db | 1472 | | |
| | | CGGAGGGCGGCGGGAGTCGCCACCAATGTTCTGTGACGAGGAGCAGCGGGGACGTGT | 1531 |
| Qy | 1261 | TCTCTGACCTTCGCGGCCCTCTGATGAAAAATGGAGCCGTGGATCAGGGGAGGTGGT | 1320 |
| Db | 1532 | | |
| | | TCTCTGACCTTCGCGGCCCTCTGATGAAAAATGGAGCCGTGGATCAGGGGAGGTGGT | 1591 |
| Qy | 1321 | TCTGAAGACCGATCTGAGGATGGGCTTCCCAAGGAATCCATCTGGATAAAAATGATGAT | 1380 |
| Db | 1592 | | |
| | | TCTGAAGACCGATCTGAGGATGGGCTTCCCAAGGAATCCATCTGGATAAAAATGATGAT | 1651 |
| Qy | 1381 | GGAGAAAAATAAAACATCTTTACATCTTTCAAGAGAGGTGATGTTATTGTGGAAGCTTTTC | 1440 |

| | | | |
|----|------|--|------|
| Db | 1652 | GGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGTGGAAAGTTTTTC | 1711 |
| Qy | 1441 | CGTTCAAAATTATTACCTCAATATTTCATCTCAGAAACGCATACAGGTGAAAAACCATACAAA | 1500 |
| Db | 1712 | CGTTCAAAATTATTACCTCAATATTTCATCTCAGAAACGCATACAGGTGAAAAACCATACAAA | 1771 |
| Qy | 1501 | TGTGAATTTTGTGGAATATGCTGCAGGCCAGAGACATCTCTGAGGTATCACTTTGGAGAGA | 1560 |
| Db | 1772 | TGTGAATTTTGTGGAATATGCTGCAGGCCAGAGACATCTCTGAGGTATCACTTTGGAGAGA | 1831 |
| Qy | 1561 | CATCACAAGGAAAAAACAACCGAGTTGTCTGCTGAAGTCAAGAACCATGTTGTAATAATCAG | 1620 |
| Db | 1832 | CATCACAAGGAAAAAACAACCGAGTTGTCTGCTGAAGTCAAGAACCATGTTGTAATAATCAG | 1891 |
| Qy | 1621 | GACACTGAAGATGCATCTATTAAACCGCTGCAGGTGCGCAAAACCAAAAAATTTGAAAAGATTT | 1680 |
| Db | 1892 | GACACTGAAGATGCATCTATTAAACCGCTGCAGGTGCGCAAAACCAAAAAATTTGAAAAGATTT | 1951 |
| Qy | 1681 | TTTGTAGTGTGCCAAAAGATGTTACAGGCAGTCCACCTGCGAAAGCAGCTTTAAGAGATGCGCT | 1740 |
| Db | 1952 | TTTGTAGTGTGCCAAAAGATGTTACAGGCAGTCCACCTGCGAAAGCAGCTTTAAGAGATGCGCT | 2011 |
| Qy | 1741 | TCCTGTTTTTTCAGAAATGTTCTGGGCAGCGTGTCTCTCACCCAGCACACAAGATACTCAG | 1800 |
| Db | 2012 | TCCTGTTTTTTCAGAAATGTTCTGGGCAGCGTGTCTCTCACCCAGCACACAAGATACTCAG | 2071 |
| Qy | 1801 | GATTTCCATAAAAATGCAGCTGATGATCAGTGTCTGATAAAGTGAATAAAAAACCTCACCCTT | 1860 |
| Db | 2072 | GATTTCCATAAAAATGCAGCTGATGATCAGTGTCTGATAAAGTGAATAAAAAACCTCACCCTT | 2131 |
| Qy | 1861 | GCTTACCTGCAACCTGTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATACCTCATC | 1920 |
| Db | 2132 | GCTTACCTGCAACCTGTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATACCTCATC | 2191 |
| Qy | 1921 | TGTAGAACCAAGCGGAGTGTACTCTCTCCGGATGGCAGTACCAACCATACCTTGAA | 1980 |
| Db | 2192 | TGTAGAACCAAGCGGAGTGTACTCTCTCCGGATGGCAGTACCAACCATACCTTGAA | 2251 |
| Qy | 1981 | GTTTAGCCCCAAGAGAACCAACCGAGACCGCAGCTGACTGCGAGATACAGGCCCAAGTGTG | 2040 |
| Db | 2252 | GTTTAGCCCCAAGAGAACCAACCGAGACCGCAGCTGACTGCGAGATACAGGCCCAAGTGTG | 2311 |
| Qy | 2041 | GATTGTCAAGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCACAATTCGCCCGCAATT | 2100 |
| Db | 2312 | GATTGTCAAGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCACAATTCGCCCGCAATT | 2371 |
| Qy | 2101 | TCCTTTGAGTAAAGTTTGATTCCAAGTATCACCTGTCCATTTTGTATCCTTCAGACATTT | 2160 |
| Db | 2372 | TCCTTTGAGTAAAGTTTGATTCCAAGTATCACCTGTCCATTTTGTATCCTTCAGACATTT | 2431 |
| Qy | 2161 | TATCCAGAAGTTTTTAATGATGCACAGAGACTGGAGCATATAATACATCTCGACGTTTCAT | 2220 |
| Db | 2432 | TATCCAGAAGTTTTTAATGATGCACAGAGACTGGAGCATATAATACATCTCGACGTTTCAT | 2491 |
| Qy | 2221 | AAAAACTGTGAAAACAAAGTCTCTGTTTGAAGTCGAGTACCGGATGCCCGCAGCGTTG | 2280 |
| Db | 2492 | AAAAACTGTGAAAACAAAGTCTCTGTTTGAAGTCGAGTACCGGATGCCCGCAGCGTTG | 2551 |
| Qy | 2281 | CTGGGAAAAAGATGTGCGCTCCCGCTCTCTAGTTTCTGTAAACCCAGGCCAAGTCTGCTTTC | 2340 |
| Db | 2552 | CTGGGAAAAAGATGTGCGCTCCCGCTCTCTAGTTTCTGTAAACCCAGGCCAAGTCTGCTTTC | 2611 |
| Qy | 2341 | CCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGCCCTCTCTGGGCCAGGC | 2400 |
| Db | 2612 | CCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGCCCTCTCTGGGCCAGGC | 2671 |
| Qy | 2401 | AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAAGTAACTTGAAGTCC | 2460 |
| Db | 2672 | AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAAGTAACTTGAAGTCC | 2731 |
| Qy | 2461 | CACAGACCAAGAGAAATGTGGGGGTCCAGGGGCCCGCCACCGGCAACAGCAATCTGAG | 2520 |

Db 2732 CACAGACCACAGAGAAATGTGGGGTCCAAAGGGCGCCACACGAGCAACAGCAATCTGAG 2791
Qy 2521 ATGTTTCTTAAACACAGTGTTCCTCCGTGACCGGTAAGACAAAAGACCCGAGACAAA 2580
Db 2792 ATGTTTCTTAAACACAGTGTTCCTCCGTGACCGGTAAGACAAAAGACCCGAGACAAA 2851
Qy 2581 TTGAAACCTCTTCCAGTAGCTCTCTCTCAGCCACCCCTCGGAGAGTAACATCAATGCT 2640
Db 2852 TTGAAACCTCTTCCAGTAGCTCTCTCTCAGCCACCCCTCGGAGAGTAACATCAATGCT 2911
Qy 2641 TCCATCGACTACCCCGCCGAAGACAGACCCCGTGGCACCTCCGGGAAGAGACTATTTC 2700
Db 2912 TCCATCGACTACCCCGCCGAAGACAGACCCCGTGGCACCTCCGGGAAGAGACTATTTC 2971
Qy 2701 TGTAACTCGAGTCCGAGCAATCTGAGCAGCAATTTGGTGAGCCCTTCCAAAAGACTG 2760
Db 2972 TGTAACTCGAGTCCGAGCAATCTGAGCAGCAATTTGGTGAGCCCTTCCAAAAGACTG 3031
Qy 2761 AAGTCCAGCGTGTGTGCTTCCAGCTTGACAGCCCGGGCCCAATTTACAGAGAGGCTAT 2820
Db 3032 AAGTCCAGCGTGTGTGCTTCCAGCTTGACAGCCCGGGCCCAATTTACAGAGAGGCTAT 3091
Qy 2821 GACCTTCCCAAGTACCATATGCTGAGAGGCATCACATCACTGTTACCGCAGCACTGTGTG 2880
Db 3092 GACCTTCCCAAGTACCATATGCTGAGAGGCATCACATCACTGTTACCGCAGCACTGTGTG 3151
Qy 2881 TATCGTGCAGCGCTGCTCCAAAACCAAGGTTCTGAGCTCCAGCGAGGTCGATCT 2940
Db 3152 TATCGTGCAGCGCTGCTCCAAAACCAAGGTTCTGAGCTCCAGCGAGGTCGATCT 3211
Qy 2941 CCAATATGCTGACTGTTTCCAGAGCCCTATGTTGGTCCCGGCCACTTTACACTTGTGTG 3000
Db 3212 CCAATATGCTGACTGTTTCCAGAGCCCTATGTTGGTCCCGGCCACTTTACACTTGTGTG 3271
Qy 3001 CCGTGTGTGCTGAGCTCCAGCTCCAGCTGCTAG----- 3034
Db 3272 CCGTGTGTGCTGAGCTCCAGCTCCAGCTGCTAGGAGGAAAGGCTGTGTATCAATCAA 3331
Qy 3035 ----- 3034
Db 3332 CACTTATCTAACAGCATGGCAACAAAGAGAAACTATGAGAAATTTTATTTGGGAATGCACAT 3391
Qy 3035 ----- 3047
Db 3392 TATGACCAATGACAAAACCTTGAATTTACTAATAGGGGAAAGAGTCTTGGTGG 3451
Qy 3048 ATGTCAGTGTCTTACTCCCATGAAATTTAAATTTTACTTCTATCTTGTGAGAGCGAATGCT 3107
Db 3452 ATGTCAGTGTCTTACTCCCATGAAATTTAAATTTTACTTCTATCTTGTGAGAGCGAATGCT 3511
Qy 3108 GAAAGCTACTGAAATGAGTGTGATGTACTGTATACATAAAACATATGAGGAATCTGCAAG 3167
Db 3512 GAAAGCTACTGAAATGAGTGTGATGTACTGTATACATAAAACATATGAGGAATCTGCAAG 3571
Qy 3168 GAACACTACAGTTGTGTAA 3186
Db 3572 GAACACTACAGTTGTGTAA 3590

RESULT 4
LOCUS CQ726295 Sequence 12229 from Patent WO02069579.
DEFINITION CQ726295 linear PAT 03-FEB-2004
ACCESSION CQ726295
VERSION CQ726295.1 GI:4228937
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE K16s, such as nucleic acid arrays, comprising a majority of

humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 12229 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
Source
1..5633
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 95.1%; Score 3029.6; DB 6; Length 5633;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 3180; Conservative 6; Mismatches 0; Indels 134; Gaps 2;
Qy 1 ATGCAATCGAAAGTGCAGGAAACATGCCAACTCAATCCCTCTTATGTACATGATGGG 60
Db 272 ATGCAATCGAAAGTGCAGGAAACATGCCAACTCAATCCCTCTTATGTACATGATGGG 331
Qy 61 CCAGAAGTGTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG 120
Db 332 CCAGAAGTGTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG 391
Qy 121 AAAGGACCGCTGTGTTCCATTCCAGCTACACAGAAAAAATGTCTCAATCCAAATCGAG 180
Db 392 AAAGGACCGCTGTGTTCCATTCCAGCTACACAGAAAAAATGTCTCAATCCAAATCGAG 451
Qy 181 GGGTATATGCCCTTGGATTGCTGTCAGCAGACCTTCCACATTCAGAAAGACCTT 240
Db 452 GGGTATATGCCCTTGGATTGCTGTCAGCAGACCTTCCACATTCAGAAAGACCTT 511
Qy 241 AATAAATGTCTTAAATGCAACACCGGCTACCTCTGTGAAACAGCAGTCTTCTGGGTT 300
Db 512 AATAAATGTCTTAAATGCAACACCGGCTACCTCTGTGAAACAGCAGTCTTCTGGGTT 571
Qy 301 GAAGCAGAGTATCTCAGTCCGCTTCAATAAAGTCAAGTCGACAGAACTCCCAAGGAA 360
Db 572 GAAGCAGAGTATCTCAGTCCGCTTCAATAAAGTCAAGTCGACAGAACTCCCAAGGAA 631
Qy 361 AAGAAATTCAGGAAATGAATTTAGCTGTGAGGTATGTGGGACAGATTTAGAGTCGCT 420
Db 632 AAGAAATTCAGGAAATGAATTTAGCTGTGAGGTATGTGGGACAGATTTAGAGTCGCT 691
Qy 421 TTTGATGTGAGATCCATGAGAAACACAAAAGATTTCTTCACTTACGGGTGTAACTG 480
Db 692 TTTGATGTGAGATCCATGAGAAACACAAAAGATTTCTTCACTTACGGGTGTAACTG 751
Qy 481 TCGGGAAGAGTTSRRSSAGCTTGGTCTTAAATTCATCGGACACATATGCG 540
Db 752 TCGGGAAGAGTTSRRSSAGCTTGGTCTTAAATTCATCGGACACATATGCG 811
Qy 541 AAATCGGGGGCCAGAAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 600
Db 812 AAATCGGGGGCCAGAAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 871
Qy 601 GAGTCTGTCAGGTGCAACCGCGCCGAGAGCATCTCTCTCTTACAAAATCTGCAATGTT 660
Db 872 GAGTCTGTCAGGTGCAACCGCGCCGAGAGCATCTCTCTCTTACAAAATCTGCAATGTT 931
Qy 661 TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTAGCAGCCGCAAGTGCACACAAA 720
Db 932 TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTAGCAGCCGCAAGTGCACACAAA 991
Qy 721 AAAACTGCTTTCCGTTACCAAGCGCGCAGACAGACTCTTCCAAAGGAGGAATGCCGTC 780
Db 992 AAAACTGCTTTCCGTTACCAAGCGCGCAGACAGACTCTTCCAAAGGAGGAATGCCGTC 1051
Qy 781 TCGAGGAGGACTTCTCTCAGTGTTCATCTTGAACCAAAATCTCAACCTGAAACGGGG 840
Db 1052 TCGAGGAGGACTTCTCTCAGTGTTCATCTTGAACCAAAATCTCAACCTGAAACGGGG 1111
Qy 841 AAGAAGCTGTGAGATGATCCCTCAGCTCGATCCGTTCCAGCTTCCAGGCTTGGCAG 900

Db 1112 AAGAAGCCTGTGATGTCATCCCTCAGCTCGATCCGTTTCAACACCTTCCAGGCTTGGCAG 1171
Qy 901 CTGGCTACCAAGAGAAAGTTGGCATTTGGCAAGAGTGAAGGAATCGGGCCNAGAGGG 960
Db 1172 CTGGCTACCAAGAGAAAGTTGGCATTTGGCAAGAGTGAAGGAATCGGGCCNAGAGAGGG 1231
Qy 961 AGCACCGACAAACGACGATTTCCAGTTTCCGAGAAGAGCTTTGGAGAAACAAATAAGGGCAGT 1020
Db 1232 AGCACCGACAAACGACGATTTCCAGTTTCCGAGAAGAGCTTTGGAGAAACAAATAAGGGCAGT 1291
Qy 1021 TGTGAGGCGCTCTCGCAAGAGAAAGAGAGTGCAGAAACATCCCAACGCGGAGAGCCCTCC 1080
Db 1292 TGTGAGGCGCTCTCGCAAGAGAAAGAGAGTGCAGAAACATCCCAACGCGGAGAGCCCTCC 1351
Qy 1081 GTGAGCGGGATCCNAGTTACCAAGTAGCAAGAGAGAGAGCCCACTCACTGCTCCGAGTGC 1140
Db 1352 GTGAGCGGGATCCNAGTTACCAAGTAGCAAGAGAGAGAGAGCCCACTCACTGCTCCGAGTGC 1411
Qy 1141 GGCAAGGCTTTAGAAACCTACCAACGAGTGGTCTTGCACTCCAGGGTCCACAGAGGAC 1200
Db 1412 GGCAAGGCTTTAGAAACCTACCAACGAGTGGTCTTGCACTCCAGGGTCCACAGAGGAC 1471
Qy 1201 CGGAGGCGCGCGGAGTGCAGCCACATGCTGTGGAAGGAGAGAGAGAGAGAGAGAGAGTGT 1260
Db 1472 CGGAGGCGCGCGGAGTGCAGCCACATGCTGTGGAAGGAGAGAGAGAGAGAGAGAGAGTGT 1531
Qy 1261 TCTCTGACCTCGCGGCGCTCTGATGAAATGGAGCGGTGGATCGAGGGAGAGTGT 1320
Db 1532 TCTCTGACCTCGCGGCGCTCTGATGAAATGGAGCGGTGGATCGAGGGAGAGTGT 1591
Qy 1321 TCTGAAGACGATCTGAGGATGGGCTTCCGAAAGGAATCCATCTGGATGAAAGATGATGAT 1380
Db 1592 TCTGAAGACGATCTGAGGATGGGCTTCCGAAAGGAATCCATCTGGATGAAAGATGATGAT 1651
Qy 1381 GGAGGAAATTAAGACATCTTACATCTTCAAGAGAGTGTAGTTATGTGGAAGTTTTC 1440
Db 1652 GGAGGAAATTAAGACATCTTACATCTTCAAGAGAGTGTAGTTATGTGGAAGTTTTC 1711
Qy 1441 GGTTCAAATTTATACCTCAATTTTCAATCTCAGACGATACAGGTGAAAGACCATACAA 1500
Db 1712 GGTTCAAATTTATACCTCAATTTTCAATCTCAGACGATACAGGTGAAAGACCATACAA 1771
Qy 1501 TGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCTGAGGTATCACTTTGAGAGA 1560
Db 1772 TGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCTGAGGTATCACTTTGAGAGA 1831
Qy 1561 CATCAAGGAAACAAACCGATGTTGCTGTGAAGTCAAGAACGATGGTAAAGATCAG 1620
Db 1832 CATCAAGGAAACAAACCGATGTTGCTGTGAAGTCAAGAACGATGGTAAAGATCAG 1891
Qy 1621 GACACTGAAGATGCACTTAAACCGCTGACAGTGCAGAACCAAAATTTGAAAGATTT 1680
Db 1892 GACACTGAAGATGCACTTAAACCGCTGACAGTGCAGAACCAAAATTTGAAAGATTT 1951
Qy 1681 TTTGATGTTGCAAGAGTTTACAGGAGTCCACCTGCAAGACGCTTAAAGAGATGCT 1740
Db 1952 TTTGATGTTGCAAGAGTTTACAGGAGTCCACCTGCAAGACGCTTAAAGAGATGCT 2011
Qy 1741 TCTGTTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCAACGACACAAAGATACAG 1800
Db 2012 TCTGTTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCAACGACACAAAGATACAG 2071
Qy 1801 GATTTCCATAAAGTGCAGTGCATGACAGTGCATGATGATGATGATGATGATGATGATGAT 1860
Db 2072 GATTTCCATAAAGTGCAGTGCATGACAGTGCATGATGATGATGATGATGATGATGATGAT 2131
Qy 1861 GCTTACCTGGACCTGTTAAAGAGATCAGCAGTGTGAACTCAGGCAAAATACCTCATC 1920
Db 2132 GCTTACCTGGACCTGTTAAAGAGATCAGCAGTGTGAACTCAGGCAAAATACCTCATC 2191
Qy 1921 TGTAGAACCAAGCGGATGTTTACTCTCTCTCGGATGGCAGTACCAACCATACCTTGAA 1980
Db 2192 TGTAGAACCAAGCGGATGTTTACTCTCTCTCGGATGGCAGTACCAACCATACCTTGAA 2251

Qy 1981 GTTAGCCCCAAGAGAAAGCAACGAGACCGCAGCTGACTGACATGACAGATACAGGCCAAGTGT 2040
Db 2252 GTTAGCCCCAAGAGAAAGCAACGAGACCGCAGCTGACTGACATGACAGATACAGGCCAAGTGT 2311
Qy 2041 GATTGTACAGAAACCTTTAAATTTATTCGTGGGGCTCTTCAAAATTTGCCCCGGCAATT 2100
Db 2312 GATTGTACAGAAACCTTTAAATTTATTCGTGGGGCTCTTCAAAATTTGCCCCGGCAATT 2371
Qy 2101 TCTTTGAGTAAAGTGTGATTTCCAAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTT 2160
Db 2372 TCTTTGAGTAAAGTGTGATTTCCAAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTT 2431
Qy 2161 TATCCAGAAATTTAAATGATGACACAGAGACTGAGAGATAAATCAATCTCTGAGCTTCA 2220
Db 2432 TATCCAGAAATTTAAATGATGACACAGAGACTGAGAGATAAATCAATCTCTGAGCTTCA 2491
Qy 2221 AAAAATGTGCAAAACAAGTCTTGTGTTAGAAAGTGCAGTACCGGATGCCCGCAGCGTTG 2280
Db 2492 AAAAATGTGCAAAACAAGTCTTGTGTTAGAAAGTGCAGTACCGGATGCCCGCAGCGTTG 2551
Qy 2281 CTGGGAAAGATGTCCTCCCTCTCTAGTTTCTGTAAACCCAAAGCCCAAGTCTGCTTTTC 2340
Db 2552 CTGGGAAAGATGTCCTCCCTCTCTAGTTTCTGTAAACCCAAAGCCCAAGTCTGCTTTTC 2611
Qy 2341 CCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGAGCCCTCTTGGGCCAGGC 2400
Db 2612 CCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGAGCCCTCTTGGGCCAGGC 2671
Qy 2401 AAGGCCCTCTGACTTCA - GGGATAGACTCTAGACATTTAGCCTTTAGCCCAAGTAACCTGAAGTC 2459
Db 2672 AAGGCCCTCTGACTTCAAGGGGATAGACTCTAGACATTTAGCCTTTAGCCCAAGTAACCTGAAGTC 2731
Qy 2460 CCAAGACCAAGCAGAGATGTGGGGTCCAAAGGGCCGACAGGCAACAGCAATCTGA 2519
Db 2732 CCAAGACCAAGCAGAGATGTGGGGTCCAAAGGGCCGACAGGCAACAGCAATCTGA 2791
Qy 2520 GATGTTTCTTAAACCAAGTGTTCCTCTGCAACCGGATGAAGACAAAGACCCGAGACAAA 2579
Db 2792 GATGTTTCTTAAACCAAGTGTTCCTCTGCAACCGGATGAAGACAAAGACCCGAGACAAA 2851
Qy 2580 ATTGAACCTCTTCCAGTAGTCTCTTTCAGCCCAACCTCGGAGCAGTAAACATCAATGG 2639
Db 2852 ATTGAACCTCTTCCAGTAGTCTCTTTCAGCCCAACCTCGGAGCAGTAAACATCAATGG 2911
Qy 2640 TTCCATCAGCTACCCCGCCAAAGACGACAGCCCGTGGCACCTCCGGAAGAGACTATTT 2699
Db 2912 TTCCATCAGCTACCCCGCCAAAGACGACAGCCCGTGGCACCTCCGGAAGAGACTATTT 2971
Qy 2700 CTGTAATCGAGTGCAGCAATACCTGCAAGCAGAAATTTGGTGAAGCCCTTCCAAAAGACT 2759
Db 2972 CTGTAATCGAGTGCAGCAATACCTGCAAGCAGAAATTTGGTGAAGCCCTTCCAAAAGACT 3031
Qy 2760 GAAGTCCAGCGTGTGCTTTCAGCTTGAACGAGCCCGGGCCAAATTAAGAGAGGCTA 2819
Db 3032 GAAGTCCAGCGTGTGCTTTCAGCTTGAACGAGCCCGGGCCAAATTAAGAGAGGCTA 3091
Qy 2820 TGACCTTCCAGTACCATATGTTGAGAGGATCACATCCTGTTTACCGAGACTGTGT 2879
Db 3092 TGACCTTCCAGTACCATATGTTGAGAGGATCACATCCTGTTTACCGAGACTGTGT 3151
Qy 2880 GTATCCGTCCGAGGCGCTCCCTCCAAACCAAGGTTTCTGAGCTCCAGGAGTGCATTC 2939
Db 3152 GTATCCGTCCGAGGCGCTCCCTCCAAACCAAGGTTTCTGAGCTCCAGGAGTGCATTC 3211
Qy 2940 TCCAAATGTGCTGACTGTTTCAAGAGCCCTATGTTGGCTCCGGGCCACTTTTACACTGTGT 2999
Db 3212 TCCAAATGTGCTGACTGTTTCAAGAGCCCTATGTTGGCTCCGGGCCACTTTTACACTGTGT 3271
Qy 3000 GCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAG - ----- 3034
Db 3272 GCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAGGAAAGGCCTGTGTATATCA 3331


```
repeat_region /note="MERSA repeat: matches 4. .183 of consensus"
9980. .10108
repeat_region /note="L12 repeat: matches 2622. .2747 of consensus"
10128. .10437
/note="AluX repeat: matches 2. .312 of consensus"
10654. .10724
/note="MIR repeat: matches 63. .137 of consensus"
10781. .11083
/note="AluX repeat: matches 1. .306 of consensus"
complement(10897. .11319)
/note="match: GSS: Em:AQ499673"
complement(10991. .11319)
/note="match: GSS: Em:AQ283579"
11462. .11759
/note="AluSg repeat: matches 1. .299 of consensus"
11878. .12182
/note="AluX repeat: matches 1. .300 of consensus"
12183. .12587
/note="L1PA8 repeat: matches 5757. .6163 of consensus"
13333. .13637
/note="MLTIG repeat: matches 3. .338 of consensus"
14201. .14507
/note="Alu repeat: matches 1. .308 of consensus"
14540. .14694
/note="L2 repeat: matches 2550. .2709 of consensus"
14764. .14868
/note="MIR repeat: matches 103. .208 of consensus"
14891. .15152
/note="AluJo repeat: matches 12. .299 of consensus"
15162. .15293
/note="L2 repeat: matches 1697. .1824 of consensus"
15317. .15544
/note="L1MB3 repeat: matches 5943. .6173 of consensus"
16306. .16603
/note="AluSc repeat: matches 1. .306 of consensus"
16610. .16901
/note="AluX repeat: matches 1. .292 of consensus"
16989. .17035
/note="MLTIC repeat: matches 420. .463 of consensus"
17036. .17338
/note="AluJo repeat: matches 1. .301 of consensus"
17339. .17730
/note="MLTIC repeat: matches 1. .420 of consensus"
17958. .18269
/note="AluX repeat: matches 1. .311 of consensus"
18288. .18315
/note="MERS1 repeat: matches 32. .59 of consensus"
18383. .18507
/note="MIR repeat: matches 23. .148 of consensus"
18544. .18650
/note="MIR repeat: matches 137. .261 of consensus"
18799. .18820
/note="11 copies 2 mer aa 100% conserved"
18835. .19184
/note="THE1A repeat: matches 1. .354 of consensus"
19231. .19522
/note="match: STS: Em:HS153YF5"
19297. .19346
/note="25 copies 2 mer ca 100% conserved"
19426. .19564
/note="MIR repeat: matches 6. .147 of consensus"
19624. .19916
/note="AluX repeat: matches 1. .295 of consensus"
20005. .20204
/note="MER20 repeat: matches 1. .218 of consensus"
20215. .20317
/note="MERSA repeat: matches 4. .108 of consensus"
20410. .20541
/note="MERSB repeat: matches 1. .120 of consensus"
20544. .20687
/note="AluSg/x repeat: matches 154. .297 of consensus"
20689. .20814
/note="FLAM_A repeat: matches 14. .136 of consensus"
```

```
repeat_region 20816. .21218
/note="Charlie4a repeat: matches 17. .427 of consensus"
repeat_region 21337. .21634
/note="AluY repeat: matches 1. .297 of consensus"

Query Match 49.0%; Score 1561.2; DB 9; Length 128871;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1482 AGGTGAAAACCATACAAATGTGAATTTTGTGAATATGTCGACGCCAGAGACATCTCT 1541
Db |||||
Qy 65718 AGGTGAAAACCATACAAATGTGAATTTTGTGAATATGTCGACGCCAGAGACATCTCT 65659
Db |||||
Qy 1542 GAGGTATCACTTGAGAGAGACATCAAAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 1601
Db |||||
Qy 65858 GAGGTATCACTTGAGAGAGACATCAAAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 65599
Db |||||
Qy 1602 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCGAAAC 1661
Db |||||
Qy 65598 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCGAAAC 65539
Db |||||
Qy 1662 CAAAAATTTGAAAAGATTTTGTGATGTCGCAAAAGATGTTACAGSCAGTCCACCTGCAAA 1721
Db |||||
Qy 65538 CAAAAATTTGAAAAGATTTTGTGATGTCGCAAAAGATGTTACAGSCAGTCCACCTGCAAA 65479
Db |||||
Qy 1722 GCAGCTTAAAGAGATGCTTCTGTTTTTTCAGAAATGTTTCTGGGSCAGCGCTGCTCTCAC 1781
Db |||||
Qy 65478 GCAGCTTAAAGAGATGCTTCTGTTTTTCAGAAATGTTTCTGGGSCAGCGCTGCTCTCAC 65419
Db |||||
Qy 1782 AGCAACAAGATGACTCAGGATTTTCATAAAAATGCAAGTATGACAGTGTGATAAAGT 1841
Db |||||
Qy 65418 AGCAACAAGATGACTCAGGATTTTCATAAAAATGCAAGTATGACAGTGTGATAAAGT 65359
Db |||||
Qy 1842 GAATAAAAACCTACCTGCTTACCTGGAGCTGTAAAAAAGAGATCAGCAGTTGAAAC 1901
Db |||||
Qy 65358 GAATAAAAACCTACCTGCTTACCTGGAGCTGTAAAAAAGAGATCAGCAGTTGAAAC 65299
Db |||||
Qy 1902 TCAGGCAAAATAACCTCATCTGTAGAAACCAAGCGCGATGTTACTCTCTCCGAGTGGCAG 1961
Db |||||
Qy 65298 TCAGGCAAAATAACCTCATCTGTAGAAACCAAGCGCGATGTTACTCTCTCCGAGTGGCAG 65239
Db |||||
Qy 1962 TACCACCCATAACCTTGAAGTTAGCCCAAGAGAGAAAGCAAGAGCCGAGAGCTGACTG 2021
Db |||||
Qy 65238 TACCACCCATAACCTTGAAGTTAGCCCAAGAGAGAAAGCAAGAGCCGAGAGCTGACTG 65179
Db |||||
Qy 2022 CAGATACAGGCCAAAGTGTGGATGTCACGAAAAACCTTTAAATTTATCCGTGGGGCTCT 2081
Db |||||
Qy 65178 CAGATACAGGCCAAAGTGTGGATGTCACGAAAAACCTTTAAATTTATCCGTGGGGCTCT 65119
Db |||||
Qy 2082 TCACAAATTCGCCGGCAATTTCTTTGAGTAAAGTTTGATTTCCAAAGTATCACCTGTCCATT 2141
Db |||||
Qy 65118 TCACAAATTCGCCGGCAATTTCTTTGAGTAAAGTTTGATTTCCAAAGTATCACCTGTCCATT 65059
Db |||||
Qy 2142 TTGTACTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACACGAGAGCTGGAGCATAA 2201
Db |||||
Qy 65058 TTGTACTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACACGAGAGCTGGAGCATAA 64999
Db |||||
Qy 2202 ATACAATCCTGACGTTTCATAAAACTGTCGAAAAAGTCTTTAGAGTTCGACGTAC 2261
Db |||||
Qy 64998 ATACAATCCTGACGTTTCATAAAACTGTCGAAAAAGTCTTTAGAGTTCGACGTAC 64939
Db |||||
Qy 2262 CGGATGCCCGCCAGCGTTGCTGGGAAAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 2321
Db |||||
Qy 64938 CGGATGCCCGCCAGCGTTGCTGGGAAAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 64879
Db |||||
Qy 2322 CAAGCCCAAGTCTGTTTCCCGGCGCAGTCAAATCCCTGCCATCTGCAAGGGAAGCA 2381
Db |||||
Qy 64878 CAAGCCCAAGTCTGTTTCCCGGCGCAGTCAAATCCCTGCCATCTGCAAGGGAAGCA 64819
Db |||||
Qy 2382 GAGCCCTCCTGGCCAGGCAAGGCCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 2441
Db |||||
Qy 64818 GAGCCCTCCTGGCCAGGCAAGGCCCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 64759
Db |||||
```

QY 2442 CCCAAGTAACCTGAAGTCCACAGACCACAGCAGAGATGTGGGGTCCCAAGGGCGCCAC 2501
| | | | |
Db 64758 CCCAAGTAACCTGAAGTCCACAGACCACAGCAGAGATGTGGGGTCCCAAGGGCGCCAC 64699
| | | | |
QY 2502 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCCAGTGTTCCTCCCTGCACCGGATAAGAC 2561
| | | | |
Db 64698 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCCAGTGTTCCTCCCTGCACCGGATAAGAC 64639
| | | | |
QY 2562 AAAAAGACCCGAGACAAAATTGAAACCTCTTCAGTAGTCTCTTCAGCCACCCCTCGG 2621
| | | | |
Db 64638 AAAAAGACCCGAGACAAAATTGAAACCTCTTCAGTAGTCTCTTCAGCCACCCCTCGG 64579
| | | | |
QY 2622 CAGCAGTAACATCAATGGTTTCATCGACTACCCCGCAAGACGACGCGGTGGGCACC 2681
| | | | |
Db 64578 CAGCAGTAACATCAATGGTTTCATCGACTACCCCGCAAGACGACGCGGTGGGCACC 64519
| | | | |
QY 2682 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATCTGCAGCAGAAATTTGGTGA 2741
| | | | |
Db 64518 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATCTGCAGCAGAAATTTGGTGA 64459
| | | | |
QY 2742 GCCCCTTCAAAAGACTGAAGTCAGCGTGTTCGCCCTTGACGTTGACCGCCGGGC 2801
| | | | |
Db 64458 GCCCCTTCAAAAGACTGAAGTCAGCGTGTTCGCCCTTGACGTTGACCGCCGGGC 64399
| | | | |
QY 2802 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGCATCACATCACT 2861
| | | | |
Db 64398 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGCATCACATCACT 64339
| | | | |
QY 2862 GTTACCGCAGGACTGTGTATTCCTCGCAGCGCGTGCCTCCCAACCAAGGTTCTCTGAG 2921
| | | | |
Db 64338 GTTACCGCAGGACTGTGTATTCCTCGCAGCGCGTGCCTCCCAACCAAGGTTCTCTGAG 64279
| | | | |
QY 2922 TTCAGCAGAGTGCATTTCCAAATGTGCTGACTGTTTCCAGAGCCCTATGFGGCTCCGG 2981
| | | | |
Db 64278 TTCAGCAGAGTGCATTTCCAAATGTGCTGACTGTTTCCAGAGCCCTATGFGGCTCCGG 64219
| | | | |
QY 2982 GCCACTTTACACTTGTGTCCTGCTGTAGTCCAGCATCCAGCTCGAGTTAGAGTCT 3041
| | | | |
Db 64218 GCCACTTTACACTTGTGTCCTGCTGTAGTCCAGCATCCAGCTCGAGTTAGAGTAT 64159
| | | | |
QY 3042 TGGTGGATGTCAGT 3055
| | | | |
Db 64158 TGCATGAGGGCGT 64145
| | | | |

RESULT 6

BD085733 10365 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085733
VERSION BD085733.1 GI:22631343
KEYWORDS JP 2001524802-A/9.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 10365)
AUTHORS Gray, J.W., Collins, C.C., Hwang, S.I., Godfrey, T., Kowbel, D. and Rommens, J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 9 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
COMMENT OS Artificial Sequence
PN JP 2001524802-A/9
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/580395, 16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11, C12Q1/68, A61K48/00
CC Description of Artificial Sequence: Genomic Sequence encoding
ZABCI

FEATURES
source
ORIGIN
Query Match 49.0%; Score 1559.6; DB 6; Length 10365;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1482 AGGTGAAACCAATACAAATGTGAATTTTGTGAATATCTGTCAGGCCCAAGACATCTCT 1541
| | | | |
Db 8286 AGGTGAAACCAATACAAATGTGAATTTTGTGAATATCTGTCAGGCCCAAGACATCTCT 8345
| | | | |
QY 1542 GAGGTATCACCTGGAGAGACATCAAGGAAACCAACCCGATGTTGCTGTAAGTCAA 1601
| | | | |
Db 8346 GAGGTATCACCTGGAGAGACATCAAGGAAACCAACCCGATGTTGCTGTAAGTCAA 8405
| | | | |
QY 1602 GAAAGATCGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGCAGAAAC 1661
| | | | |
Db 8406 GAAAGATCGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGCAGAAAC 8465
| | | | |
QY 1662 CAAAAATTTGAAAGATTTTGTGATGGTGCACAAAGATGTTACAGCAGTCCACCTGCACAA 1721
| | | | |
Db 8466 CAAAAATTTGAAAGATTTTGTGATGGTGCACAAAGATGTTACAGCAGTCCACCTGCACAA 8525
| | | | |
QY 1722 GCAGCTTAAGAGATGCTCTCTGTTTTCAGAAATGTTCTGGGCGAGCGTCTCTCTCACC 1781
| | | | |
Db 8526 GCAGCTTAAGAGATGCTCTCTGTTTTCAGAAATGTTCTGGGCGAGCGTCTCTCTCACC 8585
| | | | |
QY 1782 AGCACACAAAGATCTCAGGATTTCCATAAAAAATGCAGCTGATGACAGTGTGATAAAGT 1841
| | | | |
Db 8586 AGCACACAAAGATCTCAGGATTTCCATAAAAAATGCAGCTGATGACAGTGTGATAAAGT 8645
| | | | |
QY 1842 GAATAAAACCCCTACCCCTGCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTTGAAAC 1901
| | | | |
Db 8646 GAATAAAACCCCTACCCCTGCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTTGAAAC 8705
| | | | |
QY 1902 TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTTACTCTCTCCGGATGGCAG 1961
| | | | |
Db 8706 TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTTACTCTCTCCGGATGGCAG 8765
| | | | |
QY 1962 TACCACCATAACTTGAAGTTAGCCCAAGAGAGAAACCAAGCGAGACCGCAGCTGACTG 2021
| | | | |
Db 8766 TACCACCATAACTTGAAGTTAGCCCAAGAGAGAAACCAAGCGAGACCGCAGCTGACTG 8825
| | | | |
QY 2022 CAGATACAGGCCAAGTGTGATTTGTACGAAAAACCTTTTAAATTTATCCGTGGGGCTCT 2081
| | | | |
Db 8826 CAGATACAGGCCAAGTGTGATTTGTACGAAAAACCTTTTAAATTTATCCGTGGGGCTCT 8885
| | | | |
QY 2082 TCACAAATGCCCGCAATTTCTTTGAGTAAAGTTGATTCAGATATCACTGTCCATT 2141
| | | | |
Db 8886 TCACAAATGCCCGCAATTTCTTTGAGTAAAGTTGATTCAGATATCACTGTCCATT 8945
| | | | |
QY 2142 TTGTACCTTCAAGACATTTTATCCAGAGTTTATATGATGACACAGACACTGGAGCATAA 2201
| | | | |
Db 8946 TTGTACCTTCAAGACATTTTATCCAGAGTTTATATGATGACACAGACACTGGAGCATAA 9005
| | | | |
QY 2202 ATACAATCCTGACGTTTCATAAAAACTGTGCAAAACAAAGTCTTTGCTTAGAAGTTCGACGTAC 2261
| | | | |
Db 9006 ATACAATCCTGACGTTTCATAAAAACTGTGCAAAACAAAGTCTTTGCTTAGAAGTTCGACGTAC 9065
| | | | |
QY 2262 CGGATGCCCGCAGCGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 2321
| | | | |
Db 9066 CGGATGCCCGCAGCGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 9125
| | | | |
QY 2322 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTCGGAGAGGGAAGCA 2381
| | | | |
Db 9126 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCATCTCGGAGGGAAGCA 9185
| | | | |

| | | | | | |
|------------|---|---|------|---------------|---|
| Qy | 2382 | GAGCCCTCTCTGGCCAGGCAAGGCCCTCTGACTTCAGGATAGACTCTAGCACTTTAGC | 2441 | JOURNAL | Submitted (10-OCT-2000) Cancer Genetics, UCSF Cancer Center, 2340 |
| Db | 9186 | GAGCCCTCTCTGGCCAGGCAAGGCCCTCTGACTTCAGGATAGACTCTAGCACTTTAGC | 9245 | FEATURES | Sutter St., Rm. S151, San Francisco, CA 94706, USA |
| Qy | 2442 | CCCAAGTAACCTGAAGTCCCAAGACACAGACAGATGTGGGGTCCAGGGGCGCCAC | 2501 | source | 1. 121143 |
| Db | 9246 | CCCAAGTAACCTGAAGTCCCAAGACACAGACAGATGTGGGGTCCAGGGGCGCCAC | 9305 | | /organism="Homo sapiens" |
| Qy | 2502 | CAGCAACAGCAACTGAGATGTTTCCCTAAACAGTGTTCCTCCCTGACCGGATAGAC | 2561 | | /mol_type="genomic DNA" |
| Db | 9306 | CAGCAACAGCAACTGAGATGTTTCCCTAAACAGTGTTCCTCCCTGACCGGATAGAC | 9365 | | /db_xref="taxon:9606" |
| Qy | 2562 | AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCTCTCAGCCACCCCTCGG | 2621 | | /chromosome="20" |
| Db | 9366 | AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCTCTCAGCCACCCCTCGG | 9425 | | /map="20q13.2" |
| Qy | 2622 | CAGCAGTAACATCAATGTTTCATGACTACCCCGCCAAAGACAGACGCCGTGGGCAC | 2681 | misc_feature | /clones="97; 127; LBNL h117 and h79" |
| Db | 9426 | CAGCAGTAACATCAATGTTTCATGACTACCCCGCCAAAGACAGACGCCGTGGGCAC | 9485 | | /note="isolated in Collins lab and sequenced by the Genome |
| Qy | 2682 | TCCGGGAAGACACTATTTCTGTAATCGGAGTCCAGCAATACTGCAGCAGAAATTGGTGA | 2741 | repeat_region | complement (1. .66093) |
| Db | 9486 | TCCGGGAAGACACTATTTCTGTAATCGGAGTCCAGCAATACTGCAGCAGAAATTGGTGA | 9545 | | /note="overlaps Homo sapiens chromosome 20 clone |
| Qy | 2742 | GCCCTCTCCAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACCTTGACCGCCCGGGGC | 2801 | repeat_region | RP5-823G15 deposited in GenBank Accession Number |
| Db | 9546 | GCCCTCTCCAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACCGCCCGGGGC | 9605 | | AL354993.19" |
| Qy | 2802 | CAATACAGAGAGGCTATGACCTTCCCAAGTACATATGTCAGAGCATCATCACT | 2861 | repeat_region | 194. .468 |
| Db | 9606 | CAATACAGAGAGGCTATGACCTTCCCAAGTACATATGTCAGAGCATCATCACT | 9665 | | /rpt_family="Alu" |
| Qy | 2862 | GTTACCGCAGACTGTGTATCCGTGCGAGCGCTGCTCCCAACCAAGGTTCCCTGAG | 2921 | repeat_region | /rpt_type=dispersed |
| Db | 9666 | GTTACCGCAGACTGTGTATCCGTGCGAGCGCTGCTCCCAACCAAGGTTCCCTGAG | 9725 | misc_feature | complement (513. .653) |
| Qy | 2922 | CTCAGCGAGGTGCAATCTCCAAATGTCTGACTGTTCCAGAGCCCTATGGTGGCTCGG | 2981 | | /rpt_family="MER5" |
| Db | 9726 | CTCAGCGAGGTGCAATCTCCAAATGTCTGACTGTTCCAGAGCCCTATGGTGGCTCGG | 9785 | repeat_region | 1053. .1153 |
| Qy | 2982 | GCCACTTACACTGTGTGTCCTGCTGTGAGTCCAGCATCCAGCTCGAGTGTAGAGGTCT | 3041 | misc_feature | /rpt_type=dispersed |
| Db | 9786 | GCCACTTACACTGTGTGTCCTGCTGTGAGTCCAGCATCCAGCTCGAGTGTAGAGGTAT | 9845 | | /rpt_family="MER20" |
| Qy | 3042 | TGGTGGATGTCAGT 3055 | | repeat_region | /rpt_type=dispersed |
| Db | 9846 | TGCATGAGGGCGT 9859 | | repeat_region | 1254. .1330 |
| RESULT 7 | | | | repeat_region | /rpt_family="Alu" |
| AF312915/c | | | | repeat_region | complement (1356. .1676) |
| LOCUS | AF312915 | 121143 bp | DNA | linear | /rpt_family="Alu" |
| DEFINITION | Homo sapiens chromosome 20 clones 97 and 127, complete sequence. | | | repeat_region | 2485. .5421 |
| ACCESSION | AF312915 | | | repeat_region | /note="similar to Homo sapiens chromosome 5 clone |
| VERSION | AF312915.1 | GI:11094030 | | repeat_region | CTC-269F1 deposited in GenBank Accession Number |
| KEYWORDS | | | | repeat_region | AC011341.2" |
| SOURCE | Homo sapiens (human) | | | repeat_region | /rpt_family="Alu" |
| ORGANISM | Homo sapiens | | | repeat_region | /rpt_type=dispersed |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | repeat_region | complement (2896. .4437) |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | repeat_region | /note="similar to flow-sorted chromosome 20 HindIII |
| | 1 (bases 1 to 121143) | | | repeat_region | fragment, SC20pF17C6 deposited in GenBank Accession Number |
| | Collins, C., Volik, S., Kowbel, D., Ginzinger, D., Ylstra, B., | | | repeat_region | 294682.1" |
| | Cloutier, T., Hawkins, T., Predki, P., Martin, C., Wernick, M., | | | repeat_region | 3106. .3412 |
| | Kuo, W.L., Alberts, A. and Gray, J.W. | | | repeat_region | /rpt_family="Alu" |
| TITLE | Comprehensive genome sequence analysis of a breast cancer amplicon | | | repeat_region | /rpt_type=dispersed |
| JOURNAL | Genome Res. 11 (6), 1034-1042 (2001) | | | repeat_region | complement (5671. .5987) |
| MEDLINE | 21275464 | | | repeat_region | /rpt_family="Alu" |
| REFERENCE | 11381030 | | | repeat_region | /rpt_type=dispersed |
| AUTHORS | Volik, S., Collins, C., Gray, J., Wernick, M., Kowbel, D., Stultz, K. and | | | repeat_region | 6723. .7004 |
| | Martin, C. | | | repeat_region | /rpt_family="Alu" |
| TITLE | Direct Submission | | | repeat_region | complement (8590. .118606) |
| | | | | repeat_region | /note="similar to Mus musculus clone CT7-510C4 deposited |
| | | | | repeat_region | in GenBank Accession Number AC023610.3" |
| | | | | repeat_region | complement (9852. .10109) |
| | | | | repeat_region | /note="similar to STS WI-31010 deposited in GenBank |
| | | | | repeat_region | Accession Number G24374.1" |
| | | | | repeat_region | <10244. .>10904 |

```
/product="hypothetical"
/Note="similar to Homo sapiens hypothetical protein
OVC10-2 deposited in GenBank Accession Numbers AF230201.1,
AF000365 and in NM_018692.1"
/evidence=not experimental
repeat_region 15301..15588
/rpt_family="Alu"
/rpt_type=dispersed
misc_feature 16873..17073
/Note="similar to Homo sapiens STS WI-16697 deposited in
GenBank Accession Number G21490.1"
repeat_region 17577..17857
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(19034..19333)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 20136..20719
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 20776..21071
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(22344..22638)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 22726..22784
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(22984..23268)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(23349..23646)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(24008..24189)
/rpt_family="MIR"
/rpt_type=dispersed
repeat_region complement(24502..24795)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 24883..24941
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(25140..25425)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(25507..25805)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(26170..26352)
/rpt_family="MIR"
/rpt_type=dispersed
repeat_region 26841..27129
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 27270..33290
/rpt_family="L1"
/rpt_type=dispersed
repeat_region 28512..28730
/rpt_family="MER25"
/rpt_type=dispersed
repeat_region 34498..35009
/rpt_family="MLT2B2"
/rpt_type=dispersed
repeat_region 35035..35332
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 35343..35522
/rpt_family="Alu"
/rpt_type=dispersed
misc_feature 35593..121143
/Note="overlaps Homo sapiens clone RP4-724E16 on
```

```
chromosome 20 deposited in GenBank Accession Number
AL157838.24"
repeat_region 36137..36250
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 36366..36690
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 36949..37020
/rpt_family="MERS"
/rpt_type=dispersed
repeat_region complement(37692..37766)
/rpt_family="MERS"
/rpt_type=dispersed
repeat_region complement(38006..38276)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 38863..38984
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(39791..40072)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(40100..40395)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 40642..41118
/rpt_family="MER1"
/rpt_type=dispersed
repeat_region complement(41228..41348)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 41649..41938
Query Match 49.0%; Score 1559.6; DB 9; Length 121143;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1482 AGGTGAAACCATACAAATGTGAATTTTGTGAATATCTCTGAGCCAGAGACATCTCT 1541
Db 99470 AGGTGAAACCATACAAATGTGAATTTTGTGAATATCTCTGAGCCAGAGACATCTCT 99411
QY 1542 GAGGTATCATTGGAGAGACATCACAGGAAAAAACAACCCGATGTTGCTGAAGTCAA 1601
Db 99410 GAGGTATCATTGGAGAGACATCACAGGAAAAAACAACCCGATGTTGCTGAAGTCAA 99351
QY 1602 GAACGATGTTAAATCAGGACACTGAAGATGTCATTTAAACCGCTGACAGTGCAGCAAC 1661
Db 99350 GAACGATGTTAAATCAGGACACTGAAGATGTCATTTAAACCGCTGACAGTGCAGCAAC 99291
QY 1662 CAAAAATTTGAAAAGATTTTGTGATGGTGCCAAAGATGTTACAGCAGTCCACCTGCAGAA 1721
Db 99290 CAAAAATTTGAAAAGATTTTGTGATGGTGCCAAAGATGTTACAGCAGTCCACCTGCAGAA 99231
QY 1722 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAAATGTTCTGGGCGCGCTGCTCTCACC 1781
Db 99230 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAAATGTTCTGGGCGCGCTGCTCTCACC 99171
QY 1782 AGCACAAAGATACTCAGGATTTCCATAAAATGACAGTGCATGATGATGATGATAAAGT 1841
Db 99170 AGCACAAAGATACTCAGGATTTCCATAAAATGACAGTGCATGATGATGATAAAGT 99111
QY 1842 GAATAAAACCTTACCCCTGCTTACCTGGACCTGTATAAAAAAGAGATCAGCAGTTGAAAC 1901
Db 99110 GAATAAAACCTTACCCCTGCTTACCTGGACCTGTATAAAAAAGAGATCAGCAGTTGAAAC 99051
QY 1902 TCAGGCAATTAACCTCATCTCTAGAACCAAGCGCGATGTTACTCTCTCTCGGATGGCAG 1961
Db 99050 TCAGGCAATTAACCTCATCTCTAGAACCAAGCGCGATGTTACTCTCTCTCGGATGGCAG 98991
QY 1962 TACCACCCATTAACCTTGAAGTTAGCCCAAGAGAGAACCAAGCGAGCCGAGCTGACTG 2021
Db 98990 TACCACCCATTAACCTTGAAGTTAGCCCAAGAGAGAACCAAGCGAGCCGAGCTGACTG 98931
```

| | | | |
|----|-------|--|-------|
| Qy | 2022 | CAGATACAGCGCAAGTGTGGATTGTACGAAAAACCTTTAAATTTATCCGTGGGGCTCT | 2081 |
| Db | 98930 | CAGATACAGCGCAAGTGTGGATTGTACGAAAAACCTTTAAATTTATCCGTGGGGCTCT | 98871 |
| Qy | 2082 | TCACAAATGCCGGCAATTTCTTTGAGTAAAGTTTGATTCACAGTATCACCTGTCCATT | 2141 |
| Db | 98870 | TCACAAATGCCGGCAATTTCTTTGAGTAAAGTTTGATTCACAGTATCACCTGTCCATT | 98811 |
| Qy | 2142 | TTGTACCTTCAGACATTTTATCCAGAGTTTTAAATGATGACACAGAGCTGGAGCATAA | 2201 |
| Db | 98810 | TTGTACCTTCAGACATTTTATCCAGAGTTTTAAATGATGACACAGAGCTGGAGCATAA | 98751 |
| Qy | 2202 | ATACAATCCTGACGTTTCATAAAAACTGCGAAAAAAGTCTTGTCTTGAAGTGCACGTAC | 2261 |
| Db | 98750 | ATACAATCCTGACGTTTCATAAAAACTGCGAAAAAAGTCTTGTCTTGAAGTGCACGTAC | 98691 |
| Qy | 2262 | CGGATGCCCGCAGCGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC | 2321 |
| Db | 98690 | CGGATGCCCGCAGCGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC | 98631 |
| Qy | 2322 | CAAGCCCAAGTCTGCTTTCCGGCGCAGTCCAAATCCCTGCCATCTCGAAGGGGAGCA | 2381 |
| Db | 98630 | CAAGCCCAAGTCTGCTTTCCGGCGCAGTCCAAATCCCTGCCATCTCGAAGGGGAGCA | 98571 |
| Qy | 2382 | GAGCCCTCCTGGGCAGGCAAGGCCCTCTGCACTTCAGGGATAGACTTAGCACTTTAGC | 2441 |
| Db | 98570 | GAGCCCTCCTGGGCAGGCAAGGCCCTCTGCACTTCAGGGATAGACTTAGCACTTTAGC | 98511 |
| Qy | 2442 | CCCAAGTAACTGAAGTCCCAACAGACCAACAGAGATGTGGGGTCCAAAGGGCGCCAC | 2501 |
| Db | 98510 | CCCAAGTAACTGAAGTCCCAACAGACCAACAGAGATGTGGGGTCCAAAGGGCGCCAC | 98451 |
| Qy | 2502 | CAGGCACAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTCCCTGCACCGGATAGAC | 2561 |
| Db | 98450 | CAGGCACAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTCCCTGCACCGGATAGAC | 98391 |
| Qy | 2562 | AAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGCTCTCTCAGGCCACCCCTCGG | 2621 |
| Db | 98390 | AAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGCTCTCTCAGGCCACCCCTCGG | 98331 |
| Qy | 2622 | CAGCAGTAACTAATGTGTTCCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGCACC | 2681 |
| Db | 98330 | CAGCAGTAACTAATGTGTTCCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGCACC | 98271 |
| Qy | 2682 | TCCGGAGAGACTATTTCTGTAATCGAGTGCAGCAATACTGCGAGCAATTTGGTGA | 2741 |
| Db | 98270 | TCCGGAGAGACTATTTCTGTAATCGAGTGCAGCAATACTGCGAGCAATTTGGTGA | 98211 |
| Qy | 2742 | GCCCCCTTCCAAAAAGACTGAAGTCCAGCGTGTGGTCCCTTGACGTTGACAGCCCGGGGC | 2801 |
| Db | 98210 | GCCCCCTTCCAAAAAGACTGAAGTCCAGCGTGTGGTCCCTTGACGTTGACAGCCCGGGGC | 98151 |
| Qy | 2802 | CAATPACAGAGAGGCTATGACCTTTCCCAAGTACCATATGTCAGAGCATCATCACT | 2861 |
| Db | 98150 | CAATPACAGAGAGGCTATGACCTTTCCCAAGTACCATATGTCAGAGCATCATCACT | 98091 |
| Qy | 2862 | GTTACCGCAGGACTGTGTATCCGTGCGAGCGTGCCTCCCAACCAAGGTTCCCTGAG | 2921 |
| Db | 98090 | GTTACCGCAGGACTGTGTATCCGTGCGAGCGTGCCTCCCAACCAAGGTTCCCTGAG | 98031 |
| Qy | 2922 | CTCCAGCAGGCTCGAATTTCTCAAATGTGCTGACTGTTCCAGAGCCCTATGTTGGTCCGG | 2981 |
| Db | 98030 | CTCCAGCAGGCTCGAATTTCTCAAATGTGCTGACTGTTCCAGAGCCCTATGTTGGTCCGG | 97971 |
| Qy | 2982 | GCCACTTTACCTTGTGTCCTGCTGGTAGTCCAGCATCCAGCTCGAGTAGAGGTCT | 3041 |
| Db | 97970 | GCCACTTTACCTTGTGTCCTGCTGGTAGTCCAGCATCCAGCTCGAGTAGAGGTAT | 97911 |
| Qy | 3042 | TGGTGGATGTCAGT 3055 | |
| Db | 97910 | TGCATGAGGGGGCT 97897 | |

| | | | | | |
|-----------------------|--|--|-----------|----------------|-----------------|
| RESULT 8 | AC116668 | 105023 bp | DNA | linear | HTG 01-MAY-2002 |
| LOCUS | Trypanosoma brucei chromosome V clone RPC193-45E22, *** SEQUENCING | | | | |
| DEFINITION | IN PROGRESS ***, 3 unordered pieces. | | | | |
| ACCESSION | AC116668 | | | | |
| VERSION | AC116668.5 | GI:20376999 | | | |
| KEYWORDS | HTG; HTGS PHASE1. | | | | |
| SOURCE | Trypanosoma brucei | | | | |
| ORGANISM | Trypanosoma brucei | | | | |
| REFERENCE | Eukaryota; Euzenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma. | | | | |
| AUTHORS | 1 (bases 1 to 105023) El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K., Peterson,J., Hou,L., Zhao,H., Mason,T., Militscher,J., Pai,G., Van Aken,S., Utterback,T., Khalak,H.G., Gerard,C., Leech,V., Ullu,E., Melville,S., White,O., Adams,M.D., Donelson,J.E. and Fraser,C.M. | | | | |
| TITLE | Trypanosoma brucei GUTat10.1 RPC193-45E22 BAC genomic sequence | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | 2 (bases 1 to 105023) El-Sayed,N.M., Khalak,H. and Adams,M.D. | | | | |
| AUTHORS | Direct Submission | | | | |
| TITLE | Submitted (02-APR-2002) The Institute for Genomic Research, 9712 | | | | |
| JOURNAL | Medical Center Dr, Rockville, MD 20850, USA | | | | |
| REFERENCE | 3 (bases 1 to 105023) El-Sayed,N.M., Khalak,H. and Adams,M.D. | | | | |
| AUTHORS | Direct Submission | | | | |
| TITLE | Submitted (01-MAY-2002) The Institute for Genomic Research, 9712 | | | | |
| JOURNAL | Medical Center Dr, Rockville, MD 20850, USA | | | | |
| COMMENT | On May 1, 2002 this sequence version replaced gi:20340472. * NOTE: This is a 'working draft' sequence. It currently * consists of 3 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. | | | | |
| FEATURES | 1..105023 | | | | |
| source | /organism="Trypanosoma brucei" /mol_type="genomic DNA" /isolate="GUTat10.1" /db_xref="taxon:5691" /chromosome="V" /clone="RPC193-45E22" | | | | |
| ORIGIN | | | | | |
| Query Match | 48.9%; | Score 1556.4; | DB 2; | Length 105023; | |
| Best Local Similarity | 99.3%; | Pred. No. 0; | | | |
| Matches 1563; | Conservative 0; | Mismatches 11; | Indels 0; | Gaps 0; | |
| Qy | 1482 | AGGTGAAAAACCATCAAAATGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCT | 1541 | | |
| Db | 12569 | AGGTGAAAAACCATCAAAATGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCT | 12628 | | |
| Qy | 1542 | GAGGTATCCTTCAGGAGACATCAAGAAAAACCAACCGATGTTGCTGTAAGTCAA | 1601 | | |
| Db | 12629 | GAGGTATCCTTCAGGAGACATCAAGAAAAACCAACCGATGTTGCTGTAAGTCAA | 12688 | | |
| Qy | 1602 | GAAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCAAC | 1661 | | |
| Db | 12689 | GAAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCAAC | 12748 | | |
| Qy | 1662 | CAAAAAATTTGAAAGATTTTTTGTGTCGCAAGATGTTTACAGCAGTCCACCTGCAAA | 1721 | | |
| Db | 12749 | CAAAAAATTTGAAAGATTTTTTGTGTCGCAAGATGTTTACAGCAGTCCACCTGCAAA | 12808 | | |

```
QY 1722 GCAGCTTAAGGAGATGCGCTTCTGTTTTCAGAAATGTTCTGGCAGGCGTGTCTCTCACC 1781
Db 12809 GCAGCTTAAGGAGATGCGCTTCTGTTTTCAGAAATGTTCTGGCAGGCGTGTCTCTCACC 12868
QY 1782 AGCACACAAAGATCTCAGGATTTCCATAAAATGACGCTGATGACAGTGTCTGATAAAGT 1841
Db 12869 AGCACACAAAGATCTCAGGATTTCCATAAAATGACGCTGATGACAGTGTCTGATAAAGT 12928
QY 1842 GNAATAAAACCTACCCCTGCTTACTGGACCTGTTTAAAAAGAGATCAGCAGTTGAAAC 1901
Db 12929 GNAATAAAACCTACCCCTGCTTACTGGACCTGTTTAAAAAGAGATCAGCAGTTGAAAC 12988
QY 1902 TCAGCAAAATACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 1961
Db 12989 TCAGCAAAATACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 13048
QY 1962 TACCACCCATAACCTTGAAGTTAGCCCAAGAGAGAAAGCAAGCGAGACCGAGCTGACTG 2021
Db 13049 TACCACCCATAACCTTGAAGTTAGCCCAAGAGAGAAAGCAAGCGAGACCGAGCTGACTG 13108
QY 2022 CAGATACAGGCAAGTGTGGATGTCACGAAAAACCTTTAAATTTATCCGTGGGGCTCT 2081
Db 13109 CAGATACAGGCAAGTGTGGATGTCACGAAAAACCTTTAAATTTATCCGTGGGGCTCT 13168
QY 2082 TCACAAATGCGCGCAATTTCTTTCAGTAAAGTTTGAATCCAAAGTATCACCCTGCTCAAT 2141
Db 13169 TCACAAATGCGCGCAATTTCTTTCAGTAAAGTTTGAATCCAAAGTATCACCCTGCTCAAT 13228
QY 2142 TTGTACCTTCAAGACATTTATCCAGAAAGTTTAAATGATGACACAGAGACTGGAGCATAA 2201
Db 13229 TTGTACCTTCAAGACATTTATCCAGAAAGTTTAAATGATGACACAGAGACTGGAGCATAA 13288
QY 2202 ATACAAATCCTGAGCTTCAATAAAACCTGCGAAACCAAGTCTTGTGTTAGAAAGTCAGTAC 2261
Db 13289 ATACAAATCCTGAGCTTCAATAAAACCTGCGAAACCAAGTCTTGTGTTAGAAAGTCAGTAC 13348
QY 2262 CGGATGCCCGCAGCGTGTCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 2321
Db 13349 CGGATGCCCGCAGCGTGTCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 13408
QY 2322 CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGTCATCTCGAAAGGGGAAGCA 2381
Db 13409 CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGTCATCTCGAAAGGGGAAGCA 13468
QY 2382 GAGCCCTCTCGGCGCAGCAAGGCCCTCTGACTTTCAGGGATAGACTCTAGCAGCTTTAGC 2441
Db 13469 GAGCCCTCTCGGCGCAGCAAGGCCCTCTGACTTTCAGGGATAGACTCTAGCAGCTTTAGC 13528
QY 2442 CCCAAGTAACTGAAGTCCACAGACACAGCAGAAATGTGGGGTCCAAAGGGCGGCCAC 2501
Db 13529 CCCAAGTAACTGAAGTCCACAGACACAGCAGAAATGTGGGGTCCAAAGGGCGGCCAC 13588
QY 2502 CAGGCAACAGCAATCTGAGATGTTTCCCTAAACAGATGTTTCCCTGACCCGCGATAAGAC 2561
Db 13589 CAGGCAACAGCAATCTGAGATGTTTCCCTAAACAGATGTTTCCCTGACCCGCGATAAGAC 13648
QY 2562 AAAAGAGCCCGAGACAAATTTGAAACCTTCTCCAGTAGCTCTCTCAGCCACCCCTCGG 2621
Db 13649 AAAAGAGCCCGAGACAAATTTGAAACCTTCTCCAGTAGCTCTCTCAGCCACCCCTCGG 13708
QY 2622 CAGCGTAACATCAATGTTTCATGACTACCCCGCAGAAAGACAGACCCCGTGGGAC 2681
Db 13709 CAGCGTAACATCAATGTTTCATGACTACCCCGCAGAAAGACAGACCCCGTGGGAC 13768
QY 2682 TCCGGGAAGAGACTATTTCTGTAAATCGAGTGGCAGCAATCTGACGAGAAATTTGGTGA 2741
Db 13769 TCCGGGAAGAGACTATTTCTGTAAATCGAGTGGCAGCAATCTGACGAGAAATTTGGTGA 13828
QY 2742 GCGCCCTTCCAAAAAGACTGAAGTCAGCGTGGTGTGCGCTTGCAGCTTGACAGCCCGGGGC 2801
Db 13829 GCGCCCTTCCAAAAAGACTGAAGTCAGCGTGGTGTGCGCTTGCAGCTTGACAGCCCGGGGC 13888
```

```
QY 2802 CAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGGCATCACATCACT 2861
Db 13889 CAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGGCATCACATCACT 13948
QY 2862 GTTACCGCAGGACTGTGTGTATCCGTCGAGCGCTGCTCCAAACCAAGGTTCTGTAG 2921
Db 13949 GTTACCGCAGGACTGTGTGTATCCGTCGAGCGCTGCTCCAAACCAAGGTTCTGTAG 14008
QY 2922 CTCAGCAGAGTGTGATTTCCAAATGTCTGACTTTTCCAGAGCCCTATGTTGGTCCGG 2981
Db 14009 CTCAGCAGAGTGTGATTTCCAAATGTCTGACTTTTCCAGAGCCCTATGTTGGTCCGG 14068
QY 2982 GCCCTTTTACACTTGTGTGCTGCTGTGTAGTCCAGCATCCAGCTCGAGTTAGAGTCT 3041
Db 14069 GCCCTTTTACACTTGTGTGCTGCTGTGTAGTCCAGCATCCAGCTCGAGTTAGAGTCT 14128
QY 3042 TGGTGGATGTGAGT 3055
Db 14129 TGCATGAGGGCGT 14142

RESULT 9
LOCUS AR070326 1507 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 3 from patent US 5892010.
ACCESSION AR070326
VERSION AR070326.1 GI:7221214
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1507)
AUTHORS Gray, J., Collins, C., Hwang, S.-i., Godfrey, T., Kowbel, D. and Rommens, J.
TITLE Genes from the 20Q13 amplicon and their uses
JOURNAL Patent: US 5892010-A 3 06-APR-1999;
FEATURES
source Location/Qualifiers
1..1507
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 34.7%; Score 1106.2; DB 6; Length 1507;
Best Local Similarity 96.6%; Pred. No. 1.9e-299;
Matches 1150; Conservative 9; Mismatches 28; Indels 3; Gaps 3;
QY 1 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTCATGATGGATGG 60
Db 320 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTCATGATGGAT -GG 378
QY 61 CCAGAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 120
Db 379 CCAGAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGA -GATGCCCTTGTCAATG 437
QY 121 AAAGGGACCGCTGTTTTCATTCGGAGCTACACAGAAAAAATGTC -ATCCAAATCGA 179
Db 438 AAAGGGACCGCTGTTTTCATTCGGAGCTACACAGAAAAAATGTC -ATCCAAATCGA 497
QY 180 GGGGTATATGCCCTTGGATTCGATTCGAGCCAGACCTTTCACATTCAGAGACCT 239
Db 498 GGGGAATATGCCCTTGGATTCGATTCGAGCCAGACCTTTCACATTCAGAGACCT 557
QY 240 TAATAAATCATGCTTAATGCAACCCGCCCTTACCTCTGTGAACAGCAGATTTCTTCGGGT 299
Db 558 TAATAAATCATGCTTAATGCAACCCGCCCTTACCTCTGTGAACAGCAGATTTCTTCGGGT 617
QY 300 TGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCAGAACAGAACCTTCCCAAGGA 359
Db 618 TGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCAGAACAGAACCTTCCCAAGGA 677
QY 360 AAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGTGGGAGACATTTAGAGTCGC 419
Db 678 AAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGTGGGAGACATTTAGAGTCGC 737
```



```
Qy 420 TTTTGTGTTTGTGATCCATGAGAACACACAAAGATTCTTTTCACTTACGGGTGTAACAT 479
Db 738 TTTTGTGTTTGTGATCCATGAGAACACACAAAGATTCTTTTCACTTACGGGTGTAACAT 797

Qy 480 GTGCGGAAGAGMTTSSRSSAGCCTTGGTTCTTTTAAATAATCACATGCGGACACATAATGG 539
Db 798 GTGCGGAAGAGATTCAAGGAGCCTTGGTTCTTTTAAATAATCACATGCGGACRCATATGG 857

Qy 540 CAAATCGGGGCCGAGAGCAAACTGACAGAGCCTTGGAGATGATGTCAGCAACGATCAA 599
Db 858 CAAATCGGGGCCGAGAGCAAACTGACAGAGCCTTGGAGATGATGTCAGCAACGATCAA 917

Qy 600 CGAGTGTCTCAGGTGACCGCGCGGAGAGCATCTCTCTCTTACAAAATCTCATGGT 659
Db 918 CGAGTGTCTCAGGTGACCGCGCGGAGAGCATCTCTCTCTTGCAAAATCTCATGGT 977

Qy 660 TTGTGGCTTCTCTATTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGTGCACACAA 719
Db 978 TTGTGGCTTCTCTATTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGTGCACACAA 1037

Qy 720 AAAAACTGCTTTCGGTACGAGCGCGGACAGACTCTCCACAAGGAGGAATGCCGTC 779
Db 1038 AAAAACTGCTTTCGGTACGAGCGCGGACAGACTCTCCACAAGGAGGAATGCCGTC 1097

Qy 780 CTCGAGGGAGGACTTCTCTGCAAGTTGTTCAACTTGAGACCAAAATCTCACCTGAAACGGG 839
Db 1098 CTCGAGGGAGGACTTCTCTGCAAGTTGTTCAACTTGAGACCAAAATCTCACCTGAAACGGG 1157

Qy 840 GAAGAAGCCTGTGAGATGATCCCTCAGTCTCGATCCGTTCCACCTTCCAGGCTTGCA 899
Db 1158 GAAGAAGCCTGTGAGATGATCCCTCAGTCTCGATCCGTTCCACCTTCCAGGCTTGCA 1217

Qy 900 CTGCTGTACCAAGAAAGTTGCCATTTGCGACAGAGTGAAGAAATCGGGGCAAGAGG 959
Db 1218 KCTGCTGTACCAAGAAAGTTGCCATTTGCGACAGAGTGAAGAAATCGGGGCAAGAGG 1277

Qy 960 GAGCACGACAAACGACGATTTCGAGTTCCGGAAGAGGCTTCGAGAAACAAATAAGGGCAG 1019
Db 1278 GAGCACGACAAACGACGATTTCGAGTTCCGGAAGAGGCTTCGAGAAACAAATAAGAACCA 1337

Qy 1020 TTGTGAGCCTCTCGCAAGAAAGAGAGTGCACAACTCCACGCGCAAGCGCCCTC 1079
Db 1338 TTGTGAGCCTCTCGCAAGAAAGAGAGTGCACAACTCCACGCGCAAGCGCCCTC 1397

Qy 1080 CGTGACCGGATCCCAAGTTACCAAGTAGCAAGAGAGGCCACTCACTGCTCGAGTG 1139
Db 1398 CGTGACCGGATCCCAAGTTACCAAGTAGCAAGAGAGGCCACTCACTGCTCGAGTG 1457

Qy 1140 CGGCAAGCTTTCAGAACTTACCACGCTGGTCTTGCACTCCAGGGTCC 1189
Db 1458 CGGCAAGCTTTCAGAACTTACCACGCTGGTCTTGCACTCCAGGGTCC 1507
```

```
RESULT 10
BD085727
LOCUS BD085727 1507 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085727
VERSION BD085727.1 GI:22631337
KEYWORDS JP 2001524802-A/3.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1507)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and
Rommens,J.
```

```
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 3 04-DEC-2001:
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS Artificial Sequence
PN JP 2001524802-A/3
PD 04-DEC-2001
```

```
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11,C12Q1/68,A61K48/00
CC Description of Artificial Sequence:cc49
CC N is A, G, G, T, or U
FH Key Location/Qualifiers
FT modified base (447).
FEATURES
source
1..1507
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
```

ORIGIN

```
Query Match 34.7%; Score 1106.2; DB 6; Length 1507;
Best Local Similarity 96.6%; Pred. No. 1.9e-299;
Matches 1150; Conservative 9; Mismatches 28; Indels 3; Gaps 3;

Qy 1 ATGCAATCGAAAGTGCAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGG 60
Db 320 ATGCAATCGAAAGTGCAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGAT-GG 378

Qy 61 CCAGAAGTGTGGCAGCTCTCTTGGCAGTCCGATGGAGGATGCTTGTCAATG 120
Db 379 CCAGAAGTGTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGA-GATGCTTGTCAATG 437

Qy 121 AAAGGGACCGTGTGTTCATTCAGCTACACAAGAAAAAATGTC-ATCCAAATCGA 179
Db 438 AAAGGGCCNCTGTGTCAATTCGAGCTACACAAGAAAAAATGTCAATCCGAATCGA 497

Qy 180 GGGGTATATGCCCTTGGATTGCATGTTCTGCGAGCCAGACCTTTCACATTCAGAAACCT 239
Db 498 GGGGAATATGCCCTTGGATTGCATGTTCTGCGAGCCAGACCTTTCACATTCAGAAACCT 557

Qy 240 TAATAAACATGCTTAATGCAACACCGCCCTTACCTCTGTGAACAGCAGTTCCTCGGGT 299
Db 558 TAATAAACATGCTTAATGCAACACCGCCCTTACCTCTGTGAACAGCAGTTCCTCGGGT 617

Qy 300 TGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAAGTGCAGAACAGAACCTCCCAAGGA 359
Db 618 TGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAAGTGCAGAACAGAACCTCCCAAGGA 677

Qy 360 AAAGAATTCAGGAAATGAATTTAGCTGAGGATGTTGGGAGAGATTTAGATGCG 419
Db 678 AAAGAATTCAGGAAATGAATTTAGCTGAGGATGTTGGGAGAGATTTAGATGCG 737

Qy 420 TTTTGTGTTGTGAGATCCACATGAGAACACAAAGATTCTTTTCACTTACGGGTGTAACAT 479
Db 738 TTTTGTGTTGTGAGATCCACATGAGAACACAAAGATTCTTTTCACTTACGGGTGTAACAT 797

Qy 480 GTGCGGAAGAGMTTSSRSSAGCCTTGGTTCTTTTAAATAATCACATGCGGACACATAATGG 539
Db 798 GTGCGGAAGAGATTCAAGGAGCCTTGGTTCTTTTAAATAATCACATGCGGACRCATAATGG 857

Qy 540 CAAATCGGGGCCGAGAGCAAACTGACAGAGCCTTGGAGATGATGTCAGCAACGATCAA 599
Db 858 CAAATCGGGGCCGAGAGCAAACTGACAGAGCCTTGGAGATGATGTCAGCAACGATCAA 917

Qy 600 CGAGTGTCTCAGGTGACCGCGCGGAGAGCATCTCTCTCTTCAAAAATCTCATGGT 659
Db 918 CGAGTGTCTCAGGTGACCGCGCGGAGAGCATCTCTCTCTTGCAAAATCTCATGGT 977

Qy 660 TTGTGGCTTCTCTATTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGTGCACACAA 719
Db 978 TTGTGGCTTCTCTATTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGTGCACACAA 1037

Qy 720 AAAAACTGCTTTCGGTACGAGCGCGGAGAGCATCTCTCTCTTCCAAAGGAGGAATGCCGTC 779
Db 1038 AAAAACTGCTTTCGGTACGAGCGCGGAGAGCATCTCTCTCTTCCAAAGGAGGAATGCCGTC 1097
```


QY 780 CTCGAGGAGGACTTCCTCAGCTGTTCAACTTGAGACCAAAATCTCACCTGAAACGG 839
Db 1098 CTCGAGGAGGAGACTTCCTCAGCTGTTCAACTTGAGACCAAAATCTCACCTGAAACGG 1157
QY 840 GAAGAAGCCTGTTCAGATGATCCCTCAGCTCGATCGTTCACCACTTCACGCTTGCA 899
Db 1158 GAAGAAGCCTGTTCAGATGATCCCTCAGCTCGATCGTTCACCACTTCACGCTTGCA 1217
QY 900 GCTGCTACCAAGAAAGTTGCCATTTGCCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 959
Db 1218 KCTGGCTACCAAGAAAGTTGCCATTTGCCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1277
QY 960 GAGCACCACACGACGATTCAGTTCGAGAGGAGCTTCGAGAGGAGGAGGAGGAGGAG 1019
Db 1278 GAGCACCACACGACGATTCAGTTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1337
QY 1020 TTGTGAGGCTCTCGCAGAGAAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1079
Db 1338 TTGTGAGGCTCTCGCAGAGAAAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1397
QY 1080 COTGACGCGGATCCCAAGTTACCAAGTAGCAAGAGAGGAGGAGGAGGAGGAGGAGG 1139
Db 1398 COTGACGCGGATCCCAAGTTACCAAGTAGCAAGAGAGGAGGAGGAGGAGGAGGAGG 1457
QY 1140 CGGCAAGCTTTCAGAACCTACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1189
Db 1458 CGGCAAGCTTTCAGAACCTACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1507

RESULT 11

AL844576/c

LOCUS

DEFINITION Mouse DNA sequence from clone RP23-233M15 on chromosome 2, complete sequence.

ACCESSION AL844576

VERSION AL844576.9 GI:27848013

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 114285)

Griffiths, C.

Direct Submission

Submitted (22-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jan 22 2003 this sequence version replaced gi:27820135.

Sequence from the Mouse Genome Sequencing Consortium whole genome

shotgun may have been used to confirm this sequence. Sequence data

from the whole genome shotgun alone has only been used where it has

a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest, except on the rare

occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-233M15 is from the RPI-23 Mouse BAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACes1.6.

Location/Qualifiers

source

1..114285
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-233M15"
/clone_lib="RPI-23"

ORIGIN

Query Match 23.0%; Score 733; DB 10; Length 114285;
Best Local Similarity 70.2%; Pred. No. 5.8e-194;
Matches 1106; Conservative 0; Mismatches 415; Indels 54; Gaps 7;

QY 1482 AGGTGAAAAACCATCAAAATGTGAATTTTGTGAATATGTCGAGCCCAAGACATCTCT 1541
Db 27108 AGGTGAAAAACCATCAAAATGTGAATTTTGTGAATATGTCGAGCCCAAGACATCTCT 27049
QY 1542 GAGGTATCACTTGGAGAGACATCAAGGAAAAAACAAC---CGATGTTGCTGCTGAAGT 1598
Db 27048 GAGGTATCACTTGGAGAGACATCAAGGAAAAAACAAC---CGATGTTGCTGCTGAAGT 26989
QY 1599 CAAGAACATGTTAAATTCAGACATCAAGACATCAAGACATCAAGACATCAAGACATCA 1655
Db 26988 CAAGAACATGTTAAATTCAGACATCAAGACATCAAGACATCAAGACATCAAGACATCA 26929
QY 1656 GCAACACCAAAATTTGAAAAGATTTTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTC 1715
Db 26928 GCAACACCAAAATTTGAAAAGATTTTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTC 26869
QY 1716 TGCAAGCAGCTTAAGGAGATGCTTCTGTTTTCAGAAATGTTCTGGGCGAGCTGTCTCT 1775
Db 26868 TGCAAGCAGCTTAAGGAGATGCTTCTGTTTTCAGAAATGTTCTGGGCGAGCTGTCTCTCT 26821
QY 1776 CTCACACACACAAAGATCTCAGGATTTCCATATAAATGCAAGTATGACAGTCTGTA 1835
Db 26820 CTCACACACACAAAGATCTCAGGATTTCCATATAAATGCAAGTATGACAGTCTGTA 26764
QY 1836 TAAAGTGAATAAAACCCCTACCCCTGCTTACCTGACCTGTTTAAAGGAGATCAGCAGT 1895
Db 26763 TAAAGTGAATAAAACCCCTACCCCTGCTTACCTGACCTGTTTAAAGGAGATCAGCAGT 26707
QY 1896 TGAAACTCAGGCAATAAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCCGGA 1955
Db 26706 TGAAACTCAGGCAATAAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCCGGA 26647
QY 1956 TGCGAGTACCAACCCATTAACCTTGAAGTTAGCCCAAGAGAGAGAAACGAGACCGCAGC 2015
Db 26646 TGCGAGTACCAACCCATTAACCTTGAAGTTAGCCCAAGAGAGAGAAACGAGACCGCAGC 26614
QY 2016 TGACTGCAGATACAGGCCCAAGTGTGATGTGACGAAAAACCTTTAAATTTATCCGTGG 2075
Db 26613 TGACTGCAGATACAGGCCCAAGTGTGATGTGACGAAAAACCTTTAAATTTATCCGTGG 26554
QY 2076 GGCTCTTCACAATTTGCCCGCAATTTCTTTTGTAGTAAAGTTTGTATCCCAAGTATCACCTG 2135
Db 26553 GGCTCTTCACAATTTGCCCGCAATTTCTTTTGTAGTAAAGTTTGTATCCCAAGTATCACCTG 26494
QY 2136 TCCATTTGTACTTCAAGACATTTATTCAGAAAGTTTATGATGACACAGAGACTGA 2195
Db 26493 TCCATTTGTACTTCAAGACATTTATTCAGAAAGTTTATGATGACACAGAGACTGA 26434
QY 2196 GCATTAATACATCTCTGACGTTTCATATAAATCTGTGAAACAGTCTTGTCTTAGAGTCG 2255

```
Db 26433 GCACAGGTACAACTTGTGCTGAGCAAGTCTGTGCTGAGCAACAG 26374
Qy 2256 ACGTACCGGATGCGCGCAGAGTGTGCTGGGAAAGATGTGCTCCCTCTCTAGTTCTG 2315
Db 26373 GCGTACCGGATGCGCGTCTGCTGGGAAAGATGTGCTCCCTCTCTAGTTCTG 26314
Qy 2316 TAAACCCCAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCGCATCTGCGAAGG 2375
Db 26313 CAAGCCCAAGCCCAAGTCTGCTTTCTCACACACACTCGAAGTCCCTGCACTCAGAGAAGC 26254
Qy 2376 GAAGCAGAGCCCTCTGCGGCAGGAGAGGCCCTCTGACTTCAGGATAGACTCTAGCAC 2435
Db 26253 TCGCAGGGGCGCTCGGGGCCAAGCAAGACCCCGAGCTTCAGGACCCAGACAGCAC 26194
Qy 2436 TTTAGCCCCCAAGTAACTGAGTCCACAGACCCAGAGATGTGGGGTCCAGGGC 2495
Db 26193 TTTAGCCCCCAAGTAACTGAGTCCACAGAGTCCAACTGCTGGGGCACCAG--- 26137
Qy 2496 CGCCACAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGCACCGGA 2555
Db 26136 TGCCACAGGCGAGCAGTCTGAGTGTGTTCCCAAGGTGCGTCCCTGCTGCTATGGA 26077
Qy 2556 TAAGACAAAAGACCCGAGACAAATGAAACCTCTTCACGAGTCTCTTCTCAGCCAC 2615
Db 26076 TAAGTGAAGAGACCTGAGCCAAACTGAAGTCCCTACCGCCTCCCGTCTCAGTCCC 26017
Qy 2616 CCTGGCAGCAGTAACTCAATGTTCCATCGACTACCCGCGCAGACGACGACCCGCTG 2675
Db 26016 CCTCAGCAGTAAATAGCAACGGTTCGTTGAGTATCCCGTGAAGTGTGACGGCCCATG 25957
Qy 2676 GGCACCTCCGGGAAGAGACTATTTCTGTAATCGAGTCCAGCAATCTGCAGCAGCAAT 2735
Db 25956 GGCACAGCAAGGAGAGACTACTACTGCCATCGGAATTCGCACTGCGCAGCAGAGTA 25897
Qy 2736 TGGTACGCCCCCTTCAAAAAGACTGAAGTCAGCGTGTGCGCTTGAAGTTGACCGAC 2795
Db 25896 CAGTGAAGCAGTCCCAAGAGACTCAAGTCAGTGTGCGTGTCCCTGGACACAGAGCATGC 25837
Qy 2796 CGGGGCCAATACAGAGAGCTATGACTTCCCAAGTACCATATGTTGTCAGAGCATCAC 2855
Db 25836 AGGACCAATGGCAGAGAGGGCTTTGAGTCTCCCAAGTACCATGTTGTCAGAGCATCAC 25777
Qy 2856 ATCACTGTTACCGCAGGACTGTGTGTATCCGTGCGCAGCGCTGCTCCCAACCAAGGTT 2915
Db 25776 CTCCTGTCTACCAAGAGTGTGTGCGCCACCGCTGCTGCTGCCACAAAGCCGTTT 25717
Qy 2916 CTGAGCTCCAGCAGGTCGATTTCCAAATGTGCTGATCTGTTTCAAGAGCCCTATGTTGG 2975
Db 25716 CTGAGCCCTGGGAGGTGGAGTCAACCCAGTGTGTTGGCTGTGACAGAGCCCTACAGTGC 25657
Qy 2976 CTCGGGCGCACTTACACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3035
Db 25656 CTCTGGACCCCTGTATATCTGTGGACCCGTGGACACGACGAGGACGCCAGCCCTTGA 25597
Qy 3036 AGGCTTCTGGTGGATG 3050
Db 25596 AGGTATTACTGTAGG 25582

RESULT 12
AC023610 Mus musculus clone CT7-510C4, WORKING DRAFT SEQUENCE, 16 ordered
LOCUS AC023610
DEFINITION Mus musculus clone CT7-510C4, WORKING DRAFT SEQUENCE, 16 ordered
pieces.
AC023610
VERSION AC023610.3 GI:9256438
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 208979)
AUTHORS DOE Joint Genome Institute.
```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Sequencing of Mouse
Unpublished
2 (bases 1 to 208979)
DOE Joint Genome Institute.
Direct Submission
Submitted (16-FEB-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:8810276.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1149073
Center clone name: RG-MBAC_510C4

Summary Statistics
Consensus quality: 197269 bases at least Q40
Consensus quality: 205411 bases at least Q30
Consensus quality: 206881 bases at least Q20
Estimated insert size: 150000; pulse field gel estimation
Quality coverage: 208279; sum-of-contigs estimation
Quality coverage: 9.95 in Q20 bases; pulse field gel estimation
Quality coverage: 7.17 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 6544: contig of 6544 bp in length
* 6545: gap of unknown length
* 6645: contig of 8135 bp in length
* 14779: gap of unknown length
* 14880: contig of 1935 bp in length
* 16814: gap of unknown length
* 16915: contig of 10043 bp in length
* 26958: gap of unknown length
* 27058: contig of 11658 bp in length
* 38715: gap of unknown length
* 38816: contig of 24602 bp in length
* 63417: gap of unknown length
* 63418: contig of 16401 bp in length
* 79919: gap of unknown length
* 80019: contig of 59684 bp in length
* 139703: gap of unknown length
* 139803: contig of 2437 bp in length
* 142239: gap of unknown length
* 142340: contig of 6804 bp in length
* 149144: gap of unknown length
* 149243: contig of 5534 bp in length
* 149244: gap of unknown length
* 154778: contig of 2838 bp in length
* 154878: gap of unknown length
* 157115: contig of 3291 bp in length
* 157116: gap of unknown length
* 161107: contig of 6794 bp in length
* 161207: contig of 6794 bp in length
* 168001: gap of unknown length
* 168101: contig of 16200 bp in length
* 184301: gap of unknown length
* 184401: contig of 24579 bp in length.
Location/Qualifiers
1..208979
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="CT7-510C4"
/clone_lib="CitbC07 mouse BAC library"

FEATURES
source

ORIGIN

| Query Match | 22.7%; | Score 722.6; | DB 2; | Length 208979; |
|-----------------------|--------------|---|-----------------|--------------------|
| Best Local Similarity | 70.2%; | Prod. No. 5.6e-191; | | |
| Matches 1306; | Conservative | 0; | Mismatches 414; | Indels 55; Gaps 8; |
| Qy | 1482 | AGGTGAAAAACCATACAAATGTGAATTTTGTCAATATGTCAGACCCAGACACATCTCT | 1541 | |
| Db | 99324 | AGGTGAAAAACCATACAAATGTGAATTTCTGAGATATGCCGACGCCAAGACACATCTCT | 99383 | |
| Qy | 1542 | GAGGTATCACTTTGGAGAGACATCACAAAGGAAAAACAAAC--CGATGTTGCTGCTGAAGT | 1598 | |
| Db | 99384 | GAGGTACCACITGGAGAGACATCACAAAGACAGCAGCCGGTGATGTCGCCGCTGAGTC | 99443 | |
| Qy | 1599 | CAAGAACGATGTAATAAATTCAGGACATCAAGATGCACTATTAA---CGCTGACAGTGC | 1655 | |
| Db | 99444 | CAAAAGTGAAGCCCGAGCCAGGACCGCAGGATGCGCTACTAAAGCGCTGCTGACAGTGC | 99503 | |
| Qy | 1656 | GCAACCAAAAATTTGAAGAAGATTTTGTATGGTGCCAAAGATGTTACAGGCAGTCCACC | 1715 | |
| Db | 99504 | GCAGACCAAAAATTTAAAGAGATTTCTTGATGGTGCCAAAGATGTTAAAGGGAAGGCCACC | 99563 | |
| Qy | 1716 | TGCAAGCAGCTTAAAGGAGATGCCCTTCTGTTTTTCAGAAATGTTCTGGGCAGCGTGTCTCT | 1775 | |
| Db | 99564 | TGCCNAGCAGCTTAAAGGAGATGCCCTTCTGTCTTCAGAGTGTTCTC-----TCTC | 99611 | |
| Qy | 1776 | CTCACGACACACAAGATATCTCAGGATTTCCATAAATATGACGCTGATGACAGTGTCTGA | 1835 | |
| Db | 99612 | ACCAGCACACGACGATCTCAGGATTTCCATAAATATGACGCTGAT--AGTGTCTGA | 99668 | |
| Qy | 1836 | TAAAGTGAATAAAAACCTTACCCCTGCTTACCTGGACCTGTTAAAAAGAGATCAGCAGT | 1895 | |
| Db | 99669 | GAAGAAGGAGGAGACGCCCTGCCCTACTTATCTGGACATGACAGAGAAAGAAA---GCAGG | 99725 | |
| Qy | 1896 | TGAAACTCAGGCAAAATACCTCATCTGTAGAACCAAGCGCGATGTTTACTCTCTCCCGGA | 1955 | |
| Db | 99726 | GGAGCTCAGGCCAGCAGCCCTGTCTGCAGACTAGAGGGGGTGGGTCTTTAGCAGCGGA | 99785 | |
| Qy | 1956 | TGGCAGTACCACCCATAACCTTTGAAGTTAGCCCCAAAGAGAGAAACCGAGACCGCAGC | 2015 | |
| Db | 99786 | AGCTGG-----CCATAGGGAGAGATGATCAGGATGC | 99818 | |
| Qy | 2016 | TGACTGCAGATACAGGCCAAGTGTGGATTTGTACGAAACAACTTTAAATTTATCCGTGGG | 2075 | |
| Db | 99819 | TGACTACAGACATAAGCCGCGTGCTGACTGCCAGGACAGGCCCTTTGAATCTATCCCTGG | 99878 | |
| Qy | 2076 | GGCTCTTCACAATTTGCCCGGCAATTTCTTTGAGTAAAGTTTGATTTCCAAAGTATCACCTG | 2135 | |
| Db | 99879 | GCCGCTCACGCGCTGTCTGCAATCTCTTTGAGCAAGTGTCTGATCCCCCAGCATTTGCCCTG | 99938 | |
| Qy | 2136 | TCCATTTTGTACTTTCAAGACATTTTATCCAGAAGTTTAAATGATGCACACGAGACTGGA | 2195 | |
| Db | 99939 | CCCCTTTTGTACTTTCAAGACCTTTTATCCGGAAGTCTCTATGATGCACACAGACTTGA | 99998 | |
| Qy | 2196 | GCATAAATACAAATCTGACGTTTCATAAAACTGTGCAAAACAAGTCTTGCTTTAGAAGTGC | 2255 | |
| Db | 99999 | GCACAGGTACAACCTGACCCGACCAAGAACCGCAGCAGCAAGTCTGTGCTCAGGAAACAG | 100058 | |
| Qy | 2256 | ACGTACCGGATGCCCGCCAGCGTGTCTGGGAAAAGATGTGCCCTCCCTCTCTAGTTTCTG | 2315 | |
| Db | 100059 | CGGTACCGGGTGCCCTCCGGCTTTGCTGGGAAAAGATGTCCCTCCCTTGTCTGGGCTGCA | 100118 | |
| Qy | 2316 | TAAACCAAGCCCAAGTCTGTCTTCCGCGCGAGTCCAAATCCCTGCCATCTGCCGAAGGG | 2375 | |
| Db | 100119 | CAAGCCCAAGGCCAAGACTGCCCTTCTACCACTCGAAGTCCCTTGCACTCAGAGAGGC | 100178 | |
| Qy | 2376 | GAAGCAGAGCCCTCTCTGGGCCAAGCAAGCCCTCTGTACTTCAGGGATAGACTCTAGCAC | 2435 | |
| Db | 100179 | TCGGCAGGGGGCTTCGGGGCCAAAGCAAGCAACCCAGACTTCAGGACCAGACAACAGCAC | 100238 | |
| Qy | 2436 | TTTAGCCCCAAGTAACTTGAAGTCCCAAGACCAAGCAAGAAATGTGGGGGTCCAAGGGGC | 2495 | |
| Db | 100239 | TTTAGCCCCAAGTAACTTGAAGTCAACAGGTTCAACAGCCCAATGTGTGGGGGCACCAAG--- | 100295 | |

| QY | 2496 | CGCCACGAGCAACAGCAATCTGAGATGTTTCTTAAACACCAAGTGTTCCTCCCTGCACCGGA | 2555 |
|------------|--------|---|------------------------|
| Db | 100296 | TGCCACCAAGGAGCAGCAGTCTGAGTCTGTTTCCCAAGAGTGGCG-TCCTGCTGCTATGGA | 100354 |
| QY | 2556 | TAGACAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGTCTCTTCTCAGCCAC | 2615 |
| Db | 100355 | TARGGTGAAGAGACTGAGCCAAAACCTGAAGTCCCTACAGCTCCCGCTCTCAGTCCCC | 100414 |
| QY | 2616 | CCTCGGACGAGTAGTAACATCAATGGTTTCCATCGACTACCCCGCAAGAACGACAGCCCGTG | 2675 |
| Db | 100415 | CCTCAGCAGTAATTAATGCAACGGTTCGTGTGAGTATCCCGTGAAGGTTGACGGCCCATG | 100474 |
| QY | 2676 | GGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATACTGACAGCAGAATT | 2735 |
| Db | 100475 | GGCACAGCAAGGAGAGACTACTACTGCATCGGAATTTCTGGCAGTCCCGCAGCAGAGTA | 100534 |
| QY | 2736 | TGGTGAGCCCTTCCAAAAGACTGAAGTCCAGCGTGTTGCTTGTGAGTGTGACGACC | 2795 |
| Db | 100535 | CAGTGAGCCACATCCAAAAGACTCAAGTCCAGTCCGCTGCTCCCTGGACACAGAGCATGC | 100594 |
| QY | 2796 | CGGGGCCAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTTGTCAGAGGCATCAC | 2855 |
| Db | 100595 | AGGGACCAATGGCAGAGGGGCTTTGAGCTCCCCAAGTACCATTGTGGTCAGGAGCATCAC | 100654 |
| QY | 2856 | ATCACTGTTACCGCAGGACTGTGTATCCCTGCGCAGCGCTGCTCCCAACCAAGGTT | 2915 |
| Db | 100655 | CTCCTTGCTACCAACAGAGTGTGTGCGGCCACCGCTGTGCTGCCCCCAAGCCGCTTT | 100714 |
| QY | 2916 | CCTGAGTCCACGAGGTCGATTCTCCAAATGTGCTGACTGTTTACAGAGCCCTATGTTGG | 2975 |
| Db | 100715 | CCTGAGCCCTGGGAGGTGGAGTCAACCAAGTGTGTGCTGTGCAAGAGCCCTCAGTGC | 100774 |
| QY | 2976 | CTCCGGGCCACTTTACACTTGTGTGCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGA | 3035 |
| Db | 100775 | CTCTGGACCCCTGTATACCTGTGAGCCGTTGGACACGACGAGGAGCCGAGCCCTTGA | 100834 |
| QY | 3036 | AGGTCTTGGTGGATG 3050 | |
| Db | 100835 | AGGTATTACTGTAGG 100849 | |
| RESULT 13 | | | |
| AC118373 | | | |
| LOCUS | | 184420 bp | DNA |
| DEFINITION | | Rattus norvegicus clone CH230-356E19, *** SEQUENCING IN PROGRESS | linear HTG 19-SEP-2002 |
| | | ***, 6 unordered pieces. | |
| ACCESSION | | AC118373 | |
| VERSION | | AC118373.3 | GI:23194641 |
| KEYWORDS | | HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED. | |
| SOURCE | | Rattus norvegicus (Norway rat) | |
| ORGANISM | | Rattus norvegicus | |
| | | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| | | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; | |

RESULT 13
AC118373
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AC118373 184420 bp DNA linear HTG 19-SEP-2002
Rattus norvegicus clone CH230-356E19, *** SEQUENCING IN PROGRESS
***, 6 unordered pieces.
AC118373 GI:23194641
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 184420)
Muzny, D., Marlie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Ayanlebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cresce, A., D'Souza, L.,
David, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and separated by sized gaps filled with Ns to the estimated size, and separated may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GPZC
Center clone name: CH230-99J9
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 209040 bases at least Q40
Consensus quality: 213312 bases at least Q30
Consensus quality: 216534 bases at least Q20
Estimated insert size: 219845; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 14481: contig of 14481 bp in length
* 14482 14581: gap of unknown length
* 14582 41011: contig of 26430 bp in length
* 41012 41111: gap of unknown length
* 41112 239819: contig of 198708 bp in length
* 239820 239919: gap of unknown length
* 239920 241122: contig of 1203 bp in length
* 241123 241222: gap of unknown length
* 241223 242316: contig of 1094 bp in length
* 242317 242416: gap of unknown length
* 242417 246332: contig of 3916 bp in length.
Location/Qualifiers
1..246332
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-99J9"
1..4672
/feature="wgs_contig"
136991..139129
/feature="wgs_contig"
174114..175268
/feature="wgs_contig"
-----
ORIGIN
Query Match 22.5%; Score 716.2; DB 2; Length 246332;
Best Local Similarity 69.7%; Pred. No. 3.7e-189;
Matches 1097; Conservative 0; Mismatches 418; Indels 60; Gaps 7;
QY 1482 AGGTGAAACCATCAAAATGTAATTTGTAATATGTCGAGCCCAAGAGACATCTCT 1541
Db |||||
55734 AGGTGAAACCATCAAAATGTAATTTGTAATATGTCGAGTATGCGGAGCCCAAGAGACATCTCT 55793
QY |||||
1542 GAGGTATCATCTGGAGAGACATCAAGGAAACCA---ACCAGATGTTGCTGCTGAAGT 1598
Db |||||
55794 GAGGTATCATCTGGAGAGACATCAAGGAAACCAAGCAGCTGGGGGATGCCCTGCTGAGTC 55853
|||
```

```
QY 1599 CAAGAACGATGTTAAATAATCAGGACACTGAAGATGCACTAATTAACCCCTGACACTGGCGCA 1658
Db |||||
55854 CAAAGTGAAGCGCGAAGTCAGGAGCCGCAAGATATCTACTAACGGCTGATAGTCACA 55913
QY 1659 AACCAAAATTTGAAAGATTTTGTATGTTGCAAAAGATGTTACAGGCAGTCACCTGC 1718
Db |||||
55914 GACCAAAATTTAAAGAGATTTCTTGTGTTGTCGCAAAAGATGTTTAAAGGCGACCCACCGC 55973
QY 1719 AAAGCAGCTTAAGAGAGATGCTTCTGTTTTCAGAAATGTTCTGGGCGAGCGCTGCTCTC 1778
Db |||||
55974 CAAGCTACTTAAGAGAGATGCCCTTCTGCTTCCAGAGTGTCTGGGCGAGCAGCTGTTCTCTC 56033
QY 1779 ACCAGCAC---ACAAAGATATCTCAGGATTTCCATAAAATATCAGCTGATGATGCTGTA 1835
Db |||||
56034 ACCAGCACACAAACAGGATATCTCAGGATTTCCATAAAACGCGAGCTGATG---GTGCTGA 56090
QY 1836 TAAAGTGAATAAAACCCCTACCCCTGCTTACCTGGAGCTGTTTAAAAAGAGATCAGCAGT 1895
Db |||||
56091 GAAAGTCGAAAGAGTCTGCCCCCTGCTTATCTGGACATGCCAAGAAAGA---AAACAGG 56147
QY 1896 TGAAGCTCAGGCAATAAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCCGGA 1955
Db |||||
56148 GGATCCCGGGCCAGAGCCCTGCTGCAAACTAGAGGGGGTTGGGCCCTTAGCACGAGA 56207
QY 1956 TGGCAGTACCAACCCATTAACCTTTGAAGTTAGCCCCCAAGAGAGCAACCGGAGACCGCAGC 2015
Db |||||
56208 AGCTGG-----CCACGGGAGAGATGGATCGGAGGC 56240
QY 2016 TGACTGCAGATACAGGCCAAGTGGATGTTGTCAGCAAAACCTTTAAATTTATCCGTGGG 2075
Db |||||
56241 TGACTACAGACATAAGCCCGGTGCTGATGTCGAGGACAGGCGCTTTGAATCTATCAGTCGG 56300
QY 2076 GGCTCTTTCACAAATGCGCGCAATTTCTTTGAGTAAAGTTTGAATTCCAAGTATCACCTG 2135
Db |||||
56301 GGCACTTCAGCCCTGCTGCAATCTCTTTGAGCAAGTGTCTCATCTCCTAGCATGCTGCTG 56360
QY 2136 TCCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACCAAGAGACTGGA 2195
Db |||||
56361 CCCATTTTGTACTTTTAAACCTTTTATCCGGAAGTCTTATGATGACCAAGAGACTGA 56420
QY 2196 GCATAAATACATCTCTGACGTTTCATAAAACTGTCGAAACAGAGTCCTTGTCTAGAGTCG 2255
Db |||||
56421 GCACAGGTACAAACCCCGCAGCAGCAAGAACTTCGAGCAGCAAGTCTGTGCTGAGAGAACAG 56480
QY 2256 ACGTACCGGATGCGCGCAGCGTTGCTGGGAAAAGATGTGCTCCCTCTCTAGTTTCTG 2315
Db |||||
56481 CGGACAGGGTGCCTCCCTGCTTTGCTGGGAAGGATGTACCCCTCTGCTGCTGCTGCA 56540
QY 2316 TAAACCCCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATTCCTTGCAATCTGCGAAGGG 2375
Db |||||
56541 CAAGCCCAAGGCCAAGCCTGCTTCTCATCACAGGCTAAGGCCCTGACCCACAGAGAGAGT 56600
QY 2376 GAACGAGAGCCCTCTCTGGGCGAGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGCAC 2435
Db |||||
56601 TCGGCGAGGGGACCTTCGGGGCCAGGCAAAAGTCTCCCGACAGCTTCAGGACCGGACACAGCAC 56660
QY 2436 TTTAGCCCCCAAGTAACTCTGAAGTCCACAGACACAGAGAAATGCGGGTCCCAAGGGGC 2495
Db |||||
56661 TTTAGCCCCCAAGTAACTCTGAAGTCCACAGAGTCAAGCCAGTGCAGGGGGGCCAG--- 56717
QY 2496 CGCCACCAAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCTGACCGGA 2555
Db |||||
56718 TGCCACCAAGGCGAGCAGCAGTCAAGAGTGTTCCTTCCAAAGGTGGCGCTCCCTGCTGCTATGGA 56777
QY 2556 TAAGCAAAAAGACCCGAGACAAAATTTGAACCTTCTCCAGTAGTCTCTTCTCAGCCAC 2615
Db |||||
56778 TAAGCGGAAAAGACCTGAGCCCAAACTGAAGGCCCCCGGCTCCCCCGGCTCAGTCTCCC 56837
QY 2616 CCTCGGAGCAGTAACTCAATGGTTCATTCGACTACCCCGCCCAAGACAGCAGCCCGTG 2675
Db |||||
56838 CCTTAGCAGTGTAAATGGCAACGGTTC-----GGCTGAGTGCCCATG 56879
|||
```

```
QY 2676 GGCACCTCCGGAAGAGACTATTTCTGTATCGGAGTGCCAGCAATCTGTCAGCAGAAATT 2735
|||
Db 56880 GGGCAGCAGGAAGAGACTACTCTGTCAACGGATGTCGAGTGCCACGTCAGAAATA 56939
|||
QY 2736 TGGTGAGCCCTTCCAAAAGACTGAAGTCCAGCGGTGGTGGCCCTTGAGCTTGCACGACC 2795
|||
Db 56940 CAGTGAGCCACATCCCAAGACTCAAGTCCCTGTGGTGTCCCTGGACACAGAGCATCC 56999
|||
QY 2796 CGGGCCCAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGCTCAGAGGCATCAC 2855
|||
Db 57000 TGGGCCCAATGAGAGAGGGCTTTGAGCTCCCAAGTACCATATGCTCAGAGGCATCAC 57059
|||
QY 2856 ATCACTGTTACCGCAGGACTGTGTATCCGTCCGAGGGCTGCTCCCAAAACCAAGGTT 2915
|||
Db 57060 TTCCCTGTTACCAACAGAGTATGTGCTCACGGCTGTGCTCTCCAAAGCCCGTTT 57119
|||
QY 2916 CTGAGCTCCAGGAGGTCGATTCCTCAAAATGTGTGACTGTTCAGAGCCCTATGGTGG 2975
|||
Db 57120 CCTGAGCCCTGGGAGGTGGAGTCCGCCAGCGTGTGACTGTGCAGAGGCCCTACGGTGC 57179
|||
QY 2976 CTCGGGGCCACTTTACACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3035
|||
Db 57180 CTCTGAGCCCTGTATACCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 57239
|||
QY 3036 AGGTCTTGGTGGATG 3050
|||
Db 57240 AGGTATTACTGTAGG 57254
|||
```

```
RESULT 15
BD060160
LOCUS BD060160 469 bp DNA linear PAT 27-AUG-2002
DEFINITION Secreted expressed sequence tags (seSTs).
ACCESSION BD060160
VERSION BD060160.1 GI:22605766
KEYWORDS JP 2001518793-A/520.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 469)
REFERENCES
Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
Tracy,M., Spaulding,V. and Agostino,M.J.
Secreted expressed sequence tags (seSTs)
Patent: JP 2001518793-A 520 16-OCT-2001;
GENETICS INSTITUTE INC
COMMENT
PN JP 2001518793-A/520
PD 16-OCT-2001
PF 16-APR-1998 JP 1998543070
PR 10-APR-1997 US 08/837312
PI KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC
C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC Strandedness:
Double;
CC topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source
1..469
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
```

```
ORIGIN
Query Match 13.9%; Score 444.4; DB 6; Length 469;
Best Local Similarity 99.8%; Pred. No. 3.3e-113;
Matches 445; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1032 CTCGCAAGAAAGAGAGTGCAAACACTCCACGCGGAGCGCCCTCCGTGGACGCGGA 1091
|||
Db 19 CTAGCAAGAAAGAGAGTGCAAACACTCCACGCGGAGCGCCCTCCGTGGACGCGGA 78
|||
```

```
QY 1092 TCCCAAGTTACCCAGTAGCAAGGAGAACCCACTCACTGCTCCGAGTGGCGCAAGCTTT 1151
|||
Db 79 TCCCAAGTTACCCAGTAGCAAGGAGAACCCACTCACTGCTCCGAGTGGCGCAAGCTTT 138
|||
QY 1152 CAGAACCTTACCAACAGCTGGTCTTTCACCTCCAGGGTCCCAAGAAAGGACCGGAGGCGCG 1211
|||
Db 139 CAGAACCTTACCAACAGCTGGTCTTTCACCTCCAGGGTCCCAAGAAAGGACCGGAGGCGCG 198
|||
QY 1212 CGCGAGTCCGCCACCATGTCTGTGGACGGGAGGCGCGGGGACGTGTTCTCTGACCT 1271
|||
Db 199 CGCGAGTCCGCCACCATGTCTGTGGACGGGAGGCGCGGGGACGTGTTCTCTGACCT 258
|||
QY 1272 CGCGCCCTCTTGGATGAAATGGAGCCGTGGATCGAGGGAAGGTGTTCTTGAAGACGG 1331
|||
Db 259 CGCGCCCTCTTGGATGAAATGGAGCCGTGGATCGAGGGAAGGTGTTCTTGAAGACGG 318
|||
QY 1332 ATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGATAAAAATGATGAGGAGAAAAAT 1391
|||
Db 319 ATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGATAAAAATGATGAGGAGAAAAAT 378
|||
QY 1392 AAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAGGTTTTTCCGTTCAAATTA 1451
|||
Db 379 AAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAGGTTTTTCCGTTCAAATTA 438
|||
QY 1452 TTACCTCAATATTTCATCTCAGAACGC 1477
|||
Db 439 TTACCTCAATATTTCATCTCAGAACGC 464
|||
```

Search completed: July 4, 2005, 01:46:53
Job time : 13520.8 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 05:45:02 ; Search time 1627.82 Seconds
(without alignments)
11586.223 Million cell updates/sec

Title: US-08-731-499-10
Perfect score: 3186
Sequence: 1 ATGCAATCGAAGTGACAGG.....GGACACTACAGTTGTGTAA 3186

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-------------|--------------------|
| 1 | 3183.6 | 99.9 | 5252 | 11 ACN44987 | Acn44987 Human mrn |
| 2 | 3156.4 | 99.1 | 3183 | 2 AAV09024 | AAV09024 Homo sapi |
| 3 | 3040.6 | 95.4 | 5632 | 8 ACC50334 | ACC50334 Breast ca |
| 4 | 3040.6 | 95.4 | 5632 | 10 ADD14635 | Add14635 Human src |
| 5 | 3040.6 | 95.4 | 5632 | 13 ADR66761 | Adr66761 Human pro |
| 6 | 3040.6 | 95.4 | 5632 | 13 ADR65858 | Adr65858 Human pro |
| 7 | 3040.6 | 95.4 | 5653 | 12 ADP07267 | Adp07267 Human ZNF |
| 8 | 1561.2 | 49.0 | 14906 | 4 AAK81093 | AAK81093 Human imm |
| 9 | 1561.2 | 49.0 | 36022 | 11 ACN44986 | Acn44986 Mouse mrn |
| 10 | 1511.2 | 47.4 | 3016 | 11 ACN44985 | Acn44985 Mouse mrn |
| 11 | 1482.8 | 46.5 | 10282 | 2 AAV09023 | AAV09023 Homo sapi |
| 12 | 1106.2 | 34.7 | 1507 | 2 AAV04698 | AAV04698 Homo sapi |
| 13 | 733 | 23.0 | 26345 | 11 ACN44984 | Acn44984 Mouse gen |
| 14 | 563.6 | 17.7 | 566 | 4 AAK81095 | AAK81095 Human imm |
| 15 | 444.4 | 13.9 | 469 | 2 AAV88042 | AAV88042 EST clone |
| c 16 | 401.4 | 12.6 | 530 | 12 ACH76516 | Ach76516 Human gen |
| c 17 | 251 | 7.9 | 251 | 12 ACH90216 | Ach90216 Human gen |
| 18 | 104 | 3.3 | 5994 | 13 ACN40355 | Acn40355 Tumour-as |
| 19 | 102.4 | 3.2 | 6033 | 13 ADR25650 | Adr25650 Breast ca |
| 20 | 102.4 | 3.2 | 8156 | 6 ABT07614 | Abt07614 Human bre |

| | | | | | | | |
|---|----|------|-----|------|----|----------|--------------------|
| c | 21 | 72 | 2.3 | 424 | 3 | AAA42306 | Aaa42306 Human sec |
| | 22 | 70 | 2.2 | 2706 | 10 | ADe07836 | Ade07836 Novel cod |
| c | 23 | 66.8 | 2.1 | 1014 | 12 | ACH87715 | Ach87715 Human gen |
| c | 24 | 66.8 | 2.1 | 1229 | 12 | ACH86998 | Ach86998 Human gen |
| | 25 | 66.8 | 2.1 | 2706 | 11 | ADL71050 | Adl71050 Gene enco |
| | 26 | 66.8 | 2.1 | 2765 | 4 | AAI58185 | Aai58185 Human pol |
| | 27 | 66.8 | 2.1 | 2765 | 5 | ADQ98391 | Adq98391 DNA enco |
| | 28 | 66.8 | 2.1 | 2765 | 9 | ADB48151 | Adb48151 Novel hum |
| | 29 | 66.8 | 2.1 | 3039 | 13 | ADR25783 | Adr25783 Breast ca |
| | 30 | 66.8 | 2.1 | 3040 | 13 | ADQ87590 | Adq87590 Human tum |
| | 31 | 66.8 | 2.1 | 3617 | 4 | ABK42507 | Abk42507 Genomic s |
| | 32 | 66.8 | 2.1 | 3617 | 4 | ABK42509 | Abk42509 Genomic s |
| | 33 | 66.8 | 2.1 | 3617 | 9 | ADB60663 | Adb60663 Connectiv |
| | 34 | 66.8 | 2.1 | 3617 | 9 | ADB60665 | Adb60665 Connectiv |
| | 35 | 66.8 | 2.1 | 2909 | 11 | ADL71048 | Adl71048 Gene enco |
| | 36 | 64.2 | 2.0 | 2109 | 10 | ADJ80195 | Adj80195 Novel hum |
| | 37 | 63.8 | 2.0 | 2109 | 10 | ADJ80195 | Adj80195 Novel hum |
| | 38 | 62.8 | 2.0 | 1033 | 12 | ACH89764 | Ach89764 Human gen |
| | 39 | 62.8 | 2.0 | 2223 | 5 | AA89101 | Aas89101 DNA enco |
| | 40 | 62.8 | 2.0 | 2918 | 10 | ADB62681 | Adb62681 Human CDN |
| | 41 | 62.8 | 2.0 | 4599 | 4 | AAK90261 | Aak90261 Human dig |
| | 42 | 62.8 | 2.0 | 4599 | 4 | AAK90264 | Aak90264 Human dig |
| | 43 | 62.8 | 2.0 | 4599 | 5 | AA839893 | Aas39893 Genomic s |
| | 44 | 62.8 | 2.0 | 4599 | 5 | AA839890 | Aas39890 Genomic s |
| | 45 | 62.8 | 2.0 | 4599 | 9 | ADB32853 | Adb32853 Human nov |

ALIGNMENTS

RESULT 1
ACN44987
ID ACN44987 standard; cDNA; 5252 BP.
XX
AC ACN44987;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human mRNA sequence hCT1950762.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.
XX
OS Homo sapiens.
XX
FN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
(SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
Claim 1; SEQ ID NO 1709; Opp; English.
XX
The present invention relates to novel DNA and protein sequences which
are associated with carcinomas. The sequences are useful for: (i) for
screening drug candidates; (ii) for screening of bioactive agent capable
of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
a bioactive agent capable of modulating the activity of CAP; (iv) for
evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
and (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
determining Carcinoma Associated (CA) gene copy number. In addition, the

CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA gene
CC sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published

SQ Sequence 5252 BP; 1507 A; 1101 C; 1177 G; 1467 T; 0 U; 0 Other;

Query Match 99.9%; Score 3183.6; DB 11; Length 5252;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| Qy | 1 | ATGCAATCGAAGTGACAGGAAACATGCGCAACTCAATCCCTCTTAATGTAATGATGGG | 60 |
| Db | 25 | ATGCAATCGAAGTGACAGGAAACATGCGCAACTCAATCCCTCTTAATGTAATGATGGG | 84 |
| Qy | 61 | CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG | 120 |
| Db | 85 | CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG | 144 |
| Qy | 121 | AAAGGGACCGCTGTGTTCCATTCCGAGCTACACAAGAAAAAATGTTCATCCAAATCGAG | 180 |
| Db | 145 | AAAGGGACCGCTGTGTTCCATTCCGAGCTACACAAGAAAAAATGTTCATCCAAATCGAG | 204 |
| Qy | 181 | GGGTATATGCCCTTGGATGTGATTTCTGAGCCAGACCTTCACACATTCAGAAAGACCTT | 240 |
| Db | 205 | GGGTATATGCCCTTGGATGTGATTTCTGAGCCAGACCTTCACACATTCAGAAAGACCTT | 264 |
| Qy | 241 | AATAAATCATGTCTTAATCAACACCGGCTACCTCTGTGACAGCAGATGTTCTCGGGTT | 300 |
| Db | 265 | AATAAATCATGTCTTAATCAACACCGGCTACCTCTGTGACAGCAGATGTTCTCGGGTT | 324 |
| Qy | 301 | GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA | 360 |
| Db | 325 | GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA | 384 |
| Qy | 361 | AAGAATTCAGAGGAAATGAATTTAGCTGTGAGGTATGTGGGACAGATTTAGAGTCGCT | 420 |
| Db | 385 | AAGAATTCAGAGGAAATGAATTTAGCTGTGAGGTATGTGGGACAGATTTAGAGTCGCT | 444 |
| Qy | 421 | TTTGATGTGAGATCCATGAGAACACACAAAGATTTCTTCACTTACCGGTGTAAACATG | 480 |
| Db | 445 | TTTGATGTGAGATCCATGAGAACACACAAAGATTTCTTCACTTACCGGTGTAAACATG | 504 |
| Qy | 481 | TGCGAAGAAAGTTSSRSAGCTTTGGTTTCTTAAAAATCATGCGGACACATAATGGC | 540 |
| Db | 505 | TGCGAAGAAAGTTCAAGGACCTTGGTTTCTTAAAAATCATGCGGACACATAATGGC | 564 |
| Qy | 541 | AAATCGGGGGCCAGAAAGCAAACTGACGCAAGGCTTTGGAGATGATCCAGCAACGATCAAC | 600 |
| Db | 565 | AAATCGGGGGCCAGAAAGCAAACTGACGCAAGGCTTTGGAGATGATCCAGCAACGATCAAC | 624 |
| Qy | 601 | GAGGTGCTCCAGGTGACGCGCCGAGAGCATCTCTCTTCAAAAATCTGCATGGTT | 660 |
| Db | 625 | GAGGTGCTCCAGGTGACGCGCCGAGAGCATCTCTCTTCAAAAATCTGCATGGTT | 684 |
| Qy | 661 | TGTGGCTTCTTATTTTCAAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA | 720 |
| Db | 685 | TGTGGCTTCTTATTTTCAAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA | 744 |
| Qy | 721 | AAAACTGCTTTCGGTACGAGCAGCGCGAGACAGATCTTCCACAAAGGAGGAATGCCGTCC | 780 |
| Db | 745 | AAAACTGCTTTCGGTACGAGCAGCGCGAGACAGATCTTCCACAAAGGAGGAATGCCGTCC | 804 |
| Qy | 781 | TCGAGGAGGACTTCTCCAGTTGTTCACTTTGAGACCAAAATCTCACCTTGAACCGGG | 840 |
| Db | 805 | TCGAGGAGGACTTCTCCAGTTGTTCACTTTGAGACCAAAATCTCACCTTGAACCGGG | 864 |
| Qy | 841 | AAGAACCTGTGATGATGATCCCTCAGCTCGATCCGTTTCAACACCTTCCAGGCTTGGCAG | 900 |
| Db | 865 | AAGAACCTGTGATGATGATCCCTCAGCTCGATCCGTTTCAACACCTTCCAGGCTTGGCAG | 924 |
| Qy | 901 | CTGGCTACCAAGGAAAGTTGCCAATTTGCCAAGAGTGAAGGAATCCGGGGCAGAAAGG | 960 |

| | | | |
|----|------|---|------|
| Db | 925 | CTGGCTACCAAGGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAAGGG | 984 |
| Qy | 961 | AGCACCGACACACGACGATTCGAGTTCCGAGAAGGAGCTTGAGAGAAACAATAAGGGCAGT | 1020 |
| Db | 985 | AGCACCGACACACGACGATTCGAGTTCCGAGAAGGAGCTTGAGAGAAACAATAAGGGCAGT | 1044 |
| Qy | 1021 | TGTGACGGCTCTCTCGCAAGAGAAAGAGAAGTGCNAACACATCCCAACGCGAGGCCCTTCC | 1080 |
| Db | 1045 | TGTGACGGCTCTCTCGCAAGAGAAAGAGAAGTGCNAACACATCCCAACGCGAGGCCCTTCC | 1104 |
| Qy | 1081 | GTGACCGGATCCCAAGTTTACCCAGTAGCAAGAGAGGCCACTCACTGCTCGAGTGC | 1140 |
| Db | 1105 | GTGACCGGATCCCAAGTTTACCCAGTAGCAAGAGAGGCCACTCACTGCTCGAGTGC | 1164 |
| Qy | 1141 | GGCAAGCTTTTCAAGACCTTACCAACCTTACCCAGCTGGTCTTGACATCCAGGGTCCACNAGNAGAC | 1200 |
| Db | 1165 | GGCAAGCTTTTCAAGACCTTACCAACCTTACCCAGCTGGTCTTGACATCCAGGGTCCACNAGNAGAC | 1224 |
| Qy | 1201 | CGGAGGCGCGCGCGAGTCCGCCACCATGTCTGTGTGACCGGAGCGAGCGGGACGCTGT | 1260 |
| Db | 1225 | CGGAGGCGCGCGCGAGTCCGCCACCATGTCTGTGTGACCGGAGCGAGCGGGACGCTGT | 1284 |
| Qy | 1261 | TCTCTCACTCGCGCGCGCTCTGGATGAAATGGAGCCGTGGATCGAGGGGAAGTGGT | 1320 |
| Db | 1285 | TCTCTCACTCGCGCGCGCTCTGGATGAAATGGAGCCGTGGATCGAGGGGAAGTGGT | 1344 |
| Qy | 1321 | TCTGAACGCGATCTGAGGATGGCTTTCCCGAAGGAATCCATCTGGATTAATAATGATGAT | 1380 |
| Db | 1345 | TCTGAACGCGATCTGAGGATGGCTTTCCCGAAGGAATCCATCTGGATTAATAATGATGAT | 1404 |
| Qy | 1381 | GGAGGAAAAATAAACAATCTTACATCTTCAAGAGAGTGTATTTCTTGGAAAGTTTTTC | 1440 |
| Db | 1405 | GGAGGAAAAATAAACAATCTTACATCTTCAAGAGAGTGTATTTCTTGGAAAGTTTTTC | 1464 |
| Qy | 1441 | CGTTCAATTTATACCTCAATATTTCAATCTCAGAACGATACAGGTGAATAACCATCAAA | 1500 |
| Db | 1465 | CGTTCAATTTATACCTCAATATTTCAATCTCAGAACGATACAGGTGAATAACCATCAAA | 1524 |
| Qy | 1501 | TGTCAATTTTGTGAATATGTCGACCGCCAGAGACATCTCTGAGGTATCACTTCGAGAGA | 1560 |
| Db | 1525 | TGTGAATTTTGTGAATATGTCGACCGCCAGAGACATCTCTGAGGTATCACTTCGAGAGA | 1584 |
| Qy | 1561 | CATCAAGGAAAAAACAACCGATGTTGCTGCTGAAGTCAAGAACGATGGTAAAAATCAG | 1620 |
| Db | 1585 | CATCAAGGAAAAAACAACCGATGTTGCTGCTGAAGTCAAGAACGATGGTAAAAATCAG | 1644 |
| Qy | 1621 | GACACTGAAGATGCACTATTAAACCGCTGACAGTGGCCAAACCAAAATTTGAAAGATTT | 1680 |
| Db | 1645 | GACACTGAAGATGCACTATTAAACCGCTGACAGTGGCCAAACCAAAATTTGAAAGATTT | 1704 |
| Qy | 1681 | TTTGATGTCGCAAGATGTTTACAGGCGATCCACCTGCAAGCAGCTTAAAGGAGATGCCT | 1740 |
| Db | 1705 | TTTGATGTCGCAAGATGTTTACAGGCGATCCACCTGCAAGCAGCTTAAAGGAGATGCCT | 1764 |
| Qy | 1741 | TCTGTTTTTCAAGATGTTTGGGCGAGCTGTCTCTCAACGACACCAAAAGATCTCAG | 1800 |
| Db | 1765 | TCTGTTTTTCAAGATGTTTGGGCGAGCTGTCTCTCAACGACACCAAAAGATCTCAG | 1824 |
| Qy | 1801 | GATTTCCATAAAAATGACGCTGATGACAGTGTCTGATTAAGTGAATAAAACCTTACCCCT | 1860 |
| Db | 1825 | GATTTCCATAAAAATGACGCTGATGACAGTGTCTGATTAAGTGAATAAAACCTTACCCCT | 1884 |
| Qy | 1861 | GCTTACCTGGAACCTTTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAAATCTCATC | 1920 |
| Db | 1885 | GCTTACCTGGAACCTTTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAAATCTCATC | 1944 |
| Qy | 1921 | TGTAGAACCAAGGCGGATGTTTACTCTCTCCGGATGGCAGTACCAACCATTAACCTTGAA | 1980 |
| Db | 1945 | TGTAGAACCAAGGCGGATGTTTACTCTCTCCGGATGGCAGTACCAACCATTAACCTTGAA | 2004 |
| Qy | 1981 | GTTAGCCCCAAGAGAGCAACACCGAGACCGCAGCTGACATGCAGATACAGGCCAAGTGTG | 2040 |
| Db | 2005 | GTTAGCCCCAAGAGAGCAACACCGAGACCGCAGCTGACATGCAGATACAGGCCAAGTGTG | 2064 |

2041 GATTGTCACGAAACCTTTAAATTTATCCGTGGGGCTCTTCACAAATGCGCGCAATT 2100
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2065 GATTGTCACGAAACCTTTAAATTTATCCGTGGGGCTCTTCACAAATGCGCGCAATT 2124
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2101 TCTTTGAGTAAAGTTTGAATCCCAAGTATCACCTGTCCATTTTGACCTTCAAGACATTT 2160
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2125 TCTTTGAGTAAAGTTTGAATCCCAAGTATCACCTGTCCATTTTGACCTTCAAGACATTT 2184
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2161 TATCAGAAAGTTTAAATGATGACACAGAGACTGGAGCATAAATAACAATCCTGAGCTTCAT 2220
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2185 TATCAGAAAGTTTAAATGATGACACAGAGACTGGAGCATAAATAACAATCCTGAGCTTCAT 2244
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2221 AAAAAGTCTGAAACAGTCTTGTAGAGTCAAGTACCGATGCGGATGCGCGCAGCTTG 2280
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2245 AAAAAGTCTGAAACAGTCTTGTAGAGTCAAGTACCGATGCGGATGCGCGCAGCTTG 2304
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2281 CTGGGAAAGAGTGTCCCTCTCTAGTTTCTGTAAACCCCAAGCCCAAGTCTGCTTTC 2340
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2305 CTGGGAAAGAGTGTCCCTCTCTAGTTTCTGTAAACCCCAAGCCCAAGTCTGCTTTC 2364
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2341 CCGGCGAGTCAAAATCCCTGCTGCTGCGAAGGGGAGCAGAGCCCTCTGGGCGCAGGC 2400
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2365 CCGGCGAGTCAAAATCCCTGCTGCTGCGAAGGGGAGCAGAGCCCTCTGGGCGCAGGC 2424
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2401 AAGGCCCTCTGACTTCAAGGATAGACTCTAGCACTTTAGCCCCAAGTAACTGAAGTCC 2460
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2425 AAGGCCCTCTGACTTCAAGGATAGACTCTAGCACTTTAGCCCCAAGTAACTGAAGTCC 2484
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2461 CACAGACACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACACAGGCAACAGCAATCTGAG 2520
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2485 CACAGACACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACACAGGCAACAGCAATCTGAG 2544
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2521 ATGTTTCTTAAACACAGTGTTCCTCTGCGATAGACAAAGAGCCCGAGACAAA 2580
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2545 ATGTTTCTTAAACACAGTGTTCCTCTGCGATAGACAAAGAGCCCGAGACAAA 2604
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2581 TTGAACCTCTTCCAGTAGTCTCTCTCAGCCACCTCTGGCAGCAGTAAATCAATGGT 2640
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2605 TTGAACCTCTTCCAGTAGTCTCTCTCAGCCACCTCTGGCAGCAGTAAATCAATGGT 2664
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2641 TCCATCGAGTACCCCGCAGAAACAGCAGCCGCTGGGCACCTCCGGGAAGAGACTATTTC 2700
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2665 TCCATCGAGTACCCCGCAGAAACAGCAGCCGCTGGGCACCTCCGGGAAGAGACTATTTC 2724
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2701 TGTAAATCGAGTCCAGCAATATCGCAGCAGAAATTTGTTGAGCCCTTCCAAAAGACTG 2760
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2725 TGTAAATCGAGTCCAGCAATATCGCAGCAGAAATTTGTTGAGCCCTTCCAAAAGACTG 2784
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2761 AAGTCCAGCTGGTTGCCCTTGAAGTTCAGCCCGGGCCCAATTCAGAAAGGCTAT 2820
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2785 AAGTCCAGCTGGTTGCCCTTGAAGTTCAGCCCGGGCCCAATTCAGAAAGGCTAT 2844
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2821 GACCTTCCCAAGTACCATATGTCAGAGGCATCACATCACTGTTACCGCAGACTGTGTG 2880
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2845 GACCTTCCCAAGTACCATATGTCAGAGGCATCACATCACTGTTACCGCAGACTGTGTG 2904
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2881 TATCGTCCGAGGCGTCCCTCCCAACCAAGGTTCTGAGCTCCAGCGAGTGCATTCCT 2940
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2905 TATCGTCCGAGGCGTCCCTCCCAACCAAGGTTCTGAGCTCCAGCGAGTGCATTCCT 2964
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2941 CCAATGTGCTGACTGTTTCAGAGCCCTATGTTGCTCCGGGCCACTTTACATTTGTGTG 3000
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2965 CCAATGTGCTGACTGTTTCAGAGCCCTATGTTGCTCCGGGCCACTTTACATTTGTGTG 3024
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3001 CCTGCTGTGTCAGAGTCCAGCTCCAGCTTAGAGGCTCTGTGGATGTCAGTGCTTA 3060
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3025 CCTGCTGTGTCAGAGTCCAGCTCCAGCTTAGAGGCTCTGTGGATGTCAGTGCTTA 3084
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3061 CTCCCCATGAATTAATTTTACTTCTATCCTTTTGAAGAGCAATGGTGAAGACTACTGAA 3120
Qy ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3085 CTCCCCATGAATTAATTTTACTTCTATCCTTTGAGAGCGAATGGTGAAGACTACTGAA 3144

Qy 3121 ATAACTCTGATTTCTGTACATATAACATATGAGGAATCTGCAAGCAACACTACGTT 3180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3145 ATAACTCTGATTTCTGTACATATAACATATGAGGAATCTGCAAGCAACACTACGTT 3204
Qy 3181 GTGTAA 3186
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3205 GTGTAA 3210
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
RESULT 2
AAV09024
ID AAV09024 standard; DNA; 3183 BP.
XX AAV09024;
AC AAV09024;
XX 21-JUL-1998 (first entry)
XX Homo sapiens 20q13 amplicon ZABC-1 cDNA sequence.
DE 20q13 amplicon; chromosome 20; tumour; detection; ZABC-1 gene;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentata;
KW Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 1. 3183
FT /*tag= a
FT /product= "ZABC1 protein"
XX WO9802539-A1.
XX 22-JAN-1998.
XX 15-JUL-1997; 97WO-US012343.
XX 15-JUL-1996; 96US-00680395.
PR 16-OCT-1996; 96US-00731499.
PR 17-JAN-1997; 97US-00785532.
XX (REGC) UNIV CALIFORNIA.
XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
PI WPI; 1998-110587/10.
XX P-PSDB; AAW23975.
PT New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX Claim 1; Page 67-69; 91pp; English.
XX The sequence is that of the cDNA sequence encoding ZABC-1 (zinc finger
XX amplified in breast cancer). It maps to the core of the 20q13.2 amplicon
XX and is overexpressed in primary tumours and breast cancer cell lines
XX having 20q13.2 amplification. The sequence can be used as a probe for the
XX detection of chromosomal abnormalities at 20q13. It and other sequences
XX isolated from the 20q13 amplicon are consistently amplified in primary
XX tumours. These sequences are useful as probes or as probe targets for
XX monitoring the relative copy number of corresponding sequences from a
XX biological sample such as tumour cells. The sequences can also be used in
XX therapeutic applications for modulating the expression of the endogenous
XX gene or the activity of the gene product. Examples of therapeutic
XX approaches include antisense inhibition of gene expression, gene therapy,
XX and monoclonal antibodies that specifically bind the gene products. The
XX products can also be used in the treatment of other diseases, e.g. age-
XX related macular degeneration, Leber's congenital amaurosis and retinitis
XX pigmentata
XX Sequence 3183 BP; 925 A; 796 C; 758 G; 698 T; 0 U; 6 Other;

| | | | | | | | | | |
|---|------|---|------|--|--|--|--|--|--|
| Query Match 99.1%; Score 3156.4; DB 2; Length 3183; Best Local Similarity 99.7%; Pred. No. 0; Matches 3176; Conservative 0; Mismatches 7; Indels 3; Gaps 1; | | | | | | | | | |
| Qy | 1 | ATGCAATCGAAGTGACAGGAAAATGCAATCCCTCTTAATGATCATGGATGGG | 60 | | | | | | |
| Db | 1 | ATGCAATCGAAGTGACAGGAAAATGCAATCCCTCTTAATGATCATGGATGGG | 60 | | | | | | |
| Qy | 61 | CCAGAGATGATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGTGCCTTGTCAATG | 120 | | | | | | |
| Db | 61 | CCAGAGATGATG---GCTCTCTTGGCAGTCCGATGGAGATGGAGTGCCTTGTCAATG | 117 | | | | | | |
| Qy | 121 | AAAGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAATGTCAATCCAAATCGAG | 180 | | | | | | |
| Db | 118 | AAAGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAATGTCAATCCAAATCGAG | 177 | | | | | | |
| Qy | 181 | GGGTATATGCCCTTGGATGTGATGTTCTGCAGCCAGACCTTTCACACATTCAGAAAGCCTT | 240 | | | | | | |
| Db | 178 | CGGTATATGCCCTTGGATGTGATGTTCTGCAGCCAGACCTTTCACACATTCAGAAAGCCTT | 237 | | | | | | |
| Qy | 241 | AATTAACATGCTTAATGCAACACCGGCTTACCCTCTGTGAACGACGAGTCTTTCGGGTT | 300 | | | | | | |
| Db | 238 | AATTAACATGCTTAATGCAACACCGGCTTACCCTCTGTGAACGACGAGTCTTTCGGGTT | 297 | | | | | | |
| Qy | 301 | GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTTCCCAAGGAA | 360 | | | | | | |
| Db | 298 | GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCGAACAGAACCTTCCCAAGGAA | 357 | | | | | | |
| Qy | 361 | AAGAATTGCAAGGAAAAATGAATTTAGCTGTGAGGTATGTGGGACAGACATTTAGAGTGCCT | 420 | | | | | | |
| Db | 358 | AAGAATTGCAAGGAAAAATGAATTTAGCTGTGAGGTATGTGGGACAGACATTTAGAGTGCCT | 417 | | | | | | |
| Qy | 421 | TTTGATGTTGATGCCATGAGACACACAAGATTTCTTTCACCTTACGGGTGTAAACATG | 480 | | | | | | |
| Db | 418 | TTTGATGTTGATGCCATGAGAACACACAAGATTTCTTTCACCTTACGGGTGTAAACATG | 477 | | | | | | |
| Qy | 481 | TGCGGAAGAGMTTSSRSAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGC | 540 | | | | | | |
| Db | 478 | TGCGGAAGAGMTTSSRSAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGC | 537 | | | | | | |
| Qy | 541 | AAATCGGGGGCCAGAAGCAAACTGACGCAAGGCTTGGAGTAGTTCACGCAACGATCAAC | 600 | | | | | | |
| Db | 538 | AAATCGGGGGCCAGAAGCAAACTGACGCAAGGCTTGGAGTAGTTCACGCAACGATCAAC | 597 | | | | | | |
| Qy | 601 | GAGTCTGTCCAGGTGACGCGCCGAGAGCATCTCTCTCTTACCAATCTGATGGTT | 660 | | | | | | |
| Db | 598 | GAGTCTGTCCAGGTGACGCGCCGAGAGCATCTCTCTCTTACCAATCTGATGGTT | 657 | | | | | | |
| Qy | 661 | TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTGAGCACCGCAAGGTGCACACAAA | 720 | | | | | | |
| Db | 658 | TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTGAGCACCGCAAGGTGCACACNAA | 717 | | | | | | |
| Qy | 721 | AAAACTGCTTTTGGTACACAGAGCGCGACAGACTCTCCACAAGGAGGAATGCCGTCC | 780 | | | | | | |
| Db | 718 | AAAACTGCTTTTGGTACACAGAGCGCGACAGACTCTCCACAAGGAGGAATGCCGTCC | 777 | | | | | | |
| Qy | 781 | TCGAGGAGGACTTCTCAGTGTGTTCACTTGAGACCAAAATCTCACCTTGAAACGGG | 840 | | | | | | |
| Db | 778 | TCGAGGAGGACTTCTCAGTGTGTTCACTTGAGACCAAAATCTCACCTTGAAACGGG | 837 | | | | | | |
| Qy | 841 | AAGAAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTTCCACACCTTCCAGGCTTGGCAG | 900 | | | | | | |
| Db | 838 | AAGAAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTTCCACACCTTCCAGGCTTGGCAG | 897 | | | | | | |
| Qy | 901 | CTGGCTACCAAAGGAAAAATGTCGCAATTTGCCAAGAGTGAAGGAATCGGGGCAAGAGGG | 960 | | | | | | |
| Db | 898 | CTGGCTACCAAAGGAAAAATGTCGCAATTTGCCAAGAGTGAAGGAATCGGGGCAAGAGGG | 957 | | | | | | |
| Qy | 961 | AGCACCGACACGAGTTCGAGTTCGAGAGGAGCTTGGAGAAACAAATAAGGGCAGT | 1020 | | | | | | |
| Db | 958 | AGCACCGACACGAGTTCGAGTTCGAGAGGAGCTTGGAGAAACAAATAAGGGCAGT | 1017 | | | | | | |
| Qy | 1021 | TGTGAGGCTCTCGCAAGAGAAAGAGAGTGCRAAACACTCCACGGCGAAGCGCCTCC | 1080 | | | | | | |
| Qy | 1018 | TGTGAGGCTCTCGCAAGAGAAAGAGAGTGCRAAACACTCCACGGCGAAGCGCCTCC | 1077 | | | | | | |
| Qy | 1081 | GTGAGCGCGATCCCAAGTTTACCAGTAGCAAGAGAGAGCCCTCACTGCTCCGAGTGC | 1140 | | | | | | |
| Db | 1078 | GTGAGCGCGATCCCAAGTTTACCAGTAGCAAGAGAGAGCCCTCACTGCTCCGAGTGC | 1137 | | | | | | |
| Qy | 1141 | GGCAAGCTTTTCAAGACTTACCACCTGCTTGGATGAAAAATGGAGCCGTGATCGAGGGAAGTGGT | 1200 | | | | | | |
| Db | 1138 | GGCAAGCTTTTCAAGACTTACCACCTGCTTGGATGAAAAATGGAGCCGTGATCGAGGGAAGTGGT | 1197 | | | | | | |
| Qy | 1201 | CGAGAGCGCGCGCGAGTCCGCCACCATCTGTGTGACGCGAGGACGCGGGACGCTGT | 1260 | | | | | | |
| Db | 1198 | CGAGAGCGCGCGCGAGTCCGCCACCATCTGTGTGACGCGAGGACGCGGGACGCTGT | 1257 | | | | | | |
| Qy | 1261 | TCTCTGACCTCGCGCGCCCTCTGGATGAAAAATGGAGCCGTGATCGAGGGAAGTGGT | 1320 | | | | | | |
| Db | 1258 | TCTCTGACCTCGCGCGCCCTCTGGATGAAAAATGGAGCCGTGATCGAGGGAAGTGGT | 1317 | | | | | | |
| Qy | 1321 | TCTGAAGACGATCTGAGGATGGCTTCCCGAAGGAATCCATCTGGATGAAAAATGATGAT | 1380 | | | | | | |
| Db | 1318 | TCTGAAGACGATCTGAGGATGGCTTCCCGAAGGAATCCATCTGGATGAAAAATGATGAT | 1377 | | | | | | |
| Qy | 1381 | GGAGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTATTTATTTGTGAAAAGTTTTTC | 1440 | | | | | | |
| Db | 1378 | GGAGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTATTTATTTGTGAAAAGTTTTTC | 1437 | | | | | | |
| Qy | 1441 | CGTTCAAAATTTATACCTCAATATTCATCTCAGAACGATACAGGTGAAAAACCATACAA | 1500 | | | | | | |
| Db | 1438 | CGTTCAAAATTTATACCTCAATATTCATCTCAGAACGATACAGGTGAAAAACCATACAA | 1497 | | | | | | |
| Qy | 1501 | TGTGAATTTTGTGAATGTCTGACGCCAGAGACATCTCTGAGGTATCACTTCGAGAGA | 1560 | | | | | | |
| Db | 1498 | TGTGAATTTTGTGAATGTCTGACGCCAGAGACATCTCTGAGGTATCACTTCGAGAGA | 1557 | | | | | | |
| Qy | 1561 | CATCACAAGGAAAAAACCAGATGTTGCTGCTGAACTCAAGAACGATGTTAAATAATCAG | 1620 | | | | | | |
| Db | 1558 | CATCACAAGGAAAAAACCAGATGTTGCTGCTGAACTCAAGAACGATGTTAAATAATCAG | 1617 | | | | | | |
| Qy | 1621 | GACACTGAAGATGCATTTAAACCGCTGACAGTCCGCAACCAAAATTTGAAAAGATTT | 1680 | | | | | | |
| Db | 1618 | GACACTGAAGATGCATTTAAACCGCTGACAGTCCGCAACCAAAATTTGAAAAGATTT | 1677 | | | | | | |
| Qy | 1681 | TTTGATGCTGCAAGATGTTTACAGGAGCTCCACTGCAAAAGAGCTTTAAGAGATGCCT | 1740 | | | | | | |
| Db | 1678 | TTTGATGCTGCAAGATGTTTACAGGAGCTCCACTGCAAAAGAGCTTTAAGAGATGCCT | 1737 | | | | | | |
| Qy | 1741 | TCTGTTTTTTCAGAAATGTTCTGGGACGCTGTCTCTCAACAGCACAAAGATACTCAG | 1800 | | | | | | |
| Db | 1738 | TCTGTTTTTTCAGAAATGTTCTGGGACGCTGTCTCTCAACAGCACAAAGATACTCAG | 1797 | | | | | | |
| Qy | 1801 | GATTTCCATAAAATGACGCTGATGACAGTGTCTGATTAAGTGAATAAAACCTTACCCCT | 1860 | | | | | | |
| Db | 1798 | GATTTCCATAAAATGACGCTGATGACAGTGTCTGATTAAGTGAATAAAACCTTACCCCT | 1857 | | | | | | |
| Qy | 1861 | GCTTACCTGACCTGTTTAAAAAGAGATCAGCAGTGTGAAACTCAGGCAAAATAAATCATC | 1920 | | | | | | |
| Db | 1858 | GCTTACCTGACCTGTTTAAAAAGAGATCAGCAGTGTGAAACTCAGGCAAAATAAATCATC | 1917 | | | | | | |
| Qy | 1921 | TGTAAGAACCAAGCGGATGTTTCTCTCCCGGATGGCAGTACACCCATAAATCTTGAA | 1980 | | | | | | |
| Db | 1918 | TGTAAGAACCAAGCGGATGTTTCTCTCCCGGATGGCAGTACACCCATAAATCTTGAA | 1977 | | | | | | |
| Qy | 1981 | GTTAGCCCCCAAGAGAGCAAAACGAGACCGCAGCTGATGACATACAGGCCAAGTGTG | 2040 | | | | | | |
| Db | 1978 | GTTAGCCCCCAAGAGAGCAAAACGAGACCGCAGCTGATGACATACAGGCCAAGTGTG | 2037 | | | | | | |
| Qy | 2041 | GATTTGTCAAGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCAAAATGCGGCAATT | 2100 | | | | | | |
| Db | 2038 | GATTTGTCAAGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCAAAATGCGGCAATT | 2097 | | | | | | |
| Qy | 2101 | TCTTTGAGTAAAAAGTTGATTTCAAGTATACCTGTCTCAATTTGTACTTCAAGACATTT | 2160 | | | | | | |

| | | | | |
|----|--|------|--|------|
| Qy | | 2341 | CCGGCGCAGTCCAATCCCTGCCAATCTGCGAAGGGGAAGCAGACCCCTCTCTGGGCCAAGC | 2400 |
| Db | | 2612 | CCGGCGCAGTCCAATCCCTGCCAATCTGCGAAGGGGAAGCAGACCCCTCTCTGGGCCAAGC | 2671 |
| Qy | | 2401 | AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAAGTAACCTCAAGTCC | 2460 |
| Db | | 2672 | AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAAGTAACCTCAAGTCC | 2731 |
| Qy | | 2461 | CACAGACCACAGCAAGAATGTGGGGGTCCAGGGGC CGCACACAGGCAACAGCAATCTTGAG | 2520 |
| Db | | 2732 | CACAGACCACAGCAAGAATGTGGGGGTCCAGGGGC CGCACACAGGCAACAGCAATCTTGAG | 2791 |
| Qy | | 2521 | ATGTTTTCTTAAACACAGTGTTCCTCCCTGACACCGGATAAGACA AAAAGACCCGAGACAAA | 2580 |
| Db | | 2792 | ATGTTTTCTTAAACACAGTGTTCCTCCCTGACACCGGATAAGACA AAAAGACCCGAGACAAA | 2851 |
| Qy | | 2581 | TTGAAAACCTCTTCAGTAGTCTCTTCTCAGCCCCACCTCGGCAGCAGTAACATCAATGGT | 2640 |
| Db | | 2852 | TTGAAAACCTCTTCAGTAGTCTCTTCTCAGCCCCACCTCGGCAGCAGTAACATCAATGGT | 2911 |
| Qy | | 2641 | TCCATCGACTACCCCGCCAGAACGACAGCCC GTGGGCACCTCGGGAAGAGACTATTTC | 2700 |
| Db | | 2912 | TCCATCGACTACCCCGCCAGAACGACAGCCC GTGGGCACCTCGGGAAGAGACTATTTC | 2971 |
| Qy | | 2701 | TGTAATCGGAGTGCAGCAATACTG CAGCAGAAATTTGGTGAGCCCTTCCAA AAAAGACTG | 2760 |
| Db | | 2972 | TGTAATCGGAGTGCAGCAATACTG CAGCAGAAATTTGGTGAGCCCTTCCAA AAAAGACTG | 3031 |
| Qy | | 2761 | AAGTCCAGCGTGGTTGCCCTTGACGTTGAC CAGGCCCGGGGCCAATTACAG AAGGCTAT | 2820 |
| Db | | 3032 | AAGTCCAGCGTGGTTGCCCTTGACGTTGAC CAGGCCCGGGGCCAATTACAG AAGGCTAT | 3091 |
| Qy | | 2821 | GACCTTCCC AAGTACCATATGGTCAGAGGCATC ACATCACTGTTTACCGCAGGACTGTGTG | 2880 |
| Db | | 3092 | GACCTTCCC AAGTACCATATGGTCAGAGGCATC ACATCACTGTTTACCGCAGGACTGTGTG | 3151 |
| Qy | | 2881 | TATCCGTGCGAGGCGCTGCCCTCCCAA ACCAAGGTTCTGTAGCTCCAGCAGGTCGATPCT | 2940 |
| Db | | 3152 | TATCCGTGCGAGGCGCTGCCCTCCCAA ACCAAGGTTCTGTAGCTCCAGCAGGTCGATPCT | 3211 |
| Qy | | 2941 | CCAAATGTGCTGACTGTTT CAGAAGCCCTATG TGTGGCTCCGGGCCACATTTACCTTGTGTG | 3000 |
| Db | | 3212 | CCAAATGTGCTGACTGTTT CAGAAGCCCTATG TGTGGCTCCGGGCCACATTTACCTTGTGTG | 3271 |
| Qy | | 3001 | CTGTGCTGGTAGTCCAGCATCCAGCTCGAC GGTAG----- | 3034 |
| Db | | 3272 | CTGTGCTGGTAGTCCAGCATCCAGCTCGAC GGTAG----- | 3331 |
| Qy | | 3035 | ----- | 3034 |
| Db | | 3332 | CAC TTATCTACAGCATGGCCAAAAGAGAA ACTATGAGAA TTTATTGGGAATGCACAT | 3391 |
| Qy | | 3035 | ----- | 3047 |
| Db | | 3392 | TATCGACCAATGACAAAAAACTTGATTTCAC TTAATTAGGGGAAAAAGGCTCTGTGGTG | 3451 |
| Qy | | 3048 | ATGTCAGTGCTTACTCCCATGAAATTA AATTTTACTTCATCTCTTTGAGAAGCGAATGGT | 3107 |
| Db | | 3452 | ATGTCAGTGCTTACTCCCATGAAATTA AATTTTACTTCATCTCTTTGAGAAGCGAATGGT | 3511 |
| Qy | | 3108 | GAAGCTACTGAAATAAGCTGTGATTCTGTACTGTACATAAAAACATATCAGGAACTCTGCAAG | 3167 |
| Db | | 3512 | GAAGCTACTGAAATAAGCTGTGATTGTACTGTACATAAAAACATATCAGGAACTCTGCAAG | 3571 |
| Qy | | 3168 | GAACACTACAGTTGTGTAA 3186 | |
| Db | | 3572 | GAACACTACAGTTGTGTAA 3590 | |

RESULT 4

ADD14635

ID ADD14635 standard; cDNA; 5632 BP.

ADDI14635;
01-JAN-2004 (first entry)
Human src biomarker polynucleotide SEQ ID NO:29.
predictor set; protein tyrosine kinase activity modulator;
protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
Homo sapiens.
WO2003062395-A2.
31-JUL-2003.
17-JAN-2003; 2003WO-US001981.
18-JAN-2002; 2002US-0350061P.
(BRIM) BRISTOL-MYERS SQUIBB CO.
Huang F, Fairchild CR, Lee FY, Shaw P;
WPI; 2003-636735/60.
P-PSDB; ADD14040.
New polynucleotides and polypeptides for predicting the activity of
compounds that interact with protein tyrosine kinases and/or protein
tyrosine kinase pathways.
Claim 2; SEQ ID NO 29; 139pp; English.
The present invention describes a predictor set comprising a plurality of
polynucleotides or polypeptides whose expression pattern is predictive of
the response of cells to treatment with a compound that modulates protein
tyrosine kinase activity or members of the protein tyrosine kinase
pathway. Also described: (1) predicting whether a compound is capable of
modulating the activity of cells, comprising obtaining a sample of cells,
determining whether the cells express a plurality of markers, and
correlating the expression of the markers to the compound's ability to
modulate the activity of the cells; (2) a plurality of cell lines for
identifying polynucleotides and polypeptides whose expression levels
correlate with compound sensitivity or resistance of cells associated
with a disease state; and (3) identifying polynucleotides and
polypeptides that predict compound sensitivity or resistance of cells
associated with a disease state, comprising subjecting the plurality of
cell lines to one or more compounds, analysing the expression pattern of
polynucleotides or polypeptides that predict the sensitivity or
resistance of cells associated with a disease state by using the
expression pattern of the microarray. The polynucleotides and
polypeptides have cytostatic activities, and can be used in gene therapy.
The polynucleotides and polypeptides are useful in predicting the
activity of compounds that interact with protein tyrosine kinases and/or
protein tyrosine kinase pathways. These may be used in determining drug
sensitivity in patients to allow the development of individualized
genetic profiles which aid in treating diseases and disorders (e.g.
cancer) based on patient response at a molecular level. The present
sequence is used in the exemplification of the present invention.
Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

| | | | | | | | | | |
|----|-----------------------|--|--------------|---------------|-----|--------|--------|-------|---|
| | Query Match | 95.4% | Score | 3040.6; | DB | 10; | Length | 5632; | |
| | Best Local Similarity | 95.8%; | Pred. No. | 0; | | | | | |
| | Matches 3180; | Conservative | 6; | Mismatches | 0; | Indels | 133; | Gaps | 1 |
| Oy | 1 | ATGCAATCGAAAGTGACAGGAACAATGCCA | ACTCCCTCTTAA | TGTACATGGATGG | 60 | | | | |
| Dd | 272 | ATGCAATCGAAAGTGACAGGAACAATGCCA | ACTCCCTCTTAA | TGTACATGGATGG | 331 | | | | |
| Oy | 61 | CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGCAGAGTCCTTTGTCAATG | 120 | | | | | | |

Db 332 |||||CCAGAAAGTGA|||TGGCAGCTCTCTTGGCAGTCCGATGGAGATGCGGTGTCATG 391
Qy 121 AAAGGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTCATCCAATCGAG 180
Db 392 AAAGGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTCATCCAATCGAG 451
Qy 181 GGGTATATGCCCTTGGATTTGATGTTCTGCGAGCCAGACCTTCACACATTCAGAAAGACCTT 240
Db 452 GGGTATATGCCCTTGGATTTGATGTTCTGCGAGCCAGACCTTCACACATTCAGAAAGACCTT 511
Qy 241 AATAAACATGCTTAAATGCAACACCGGCTTACCCTCTGTGTGAACGACAGTTCTTCGGGTT 300
Db 512 AATAAACATGCTTAAATGCAACACCGGCTTACCCTCTGTGAACGACAGTTCTTCGGGTT 571
Qy 301 GAACGAGATATCTCAGTCCGCTTGATAAAGTCAAGTGCAGACAGTCCCAAGAA 360
Db 572 GAACGAGATATCTCAGTCCGCTTGATAAAGTCAAGTGCAGACAGTCCCAAGAA 631
Qy 361 AAGAAATCGAAGGAAAAATGAATTTAGCTGTGAGGTATGTGGGACACATTTAGAGTCGCT 420
Db 632 AAGAAATTCGAAGGAAAAATGAATTTAGCTGTGAGGTATGTGGGACACATTTAGAGTCGCT 691
Qy 421 TTTGATGTTGAGATCCACATGAGAAACACAAAAGATTTCTTCACTTACGGGTGTAACATG 480
Db 692 TTTGATGTTGAGATCCACATGAGAAACACAAAAGATTTCTTCACTTACGGGTGTAACATG 751
Qy 481 TGCGGAAGAAAGTTSRRSSAGCCTGTGTTCTTAAATAATCACATGCGGACACATAATGGC 540
Db 752 TGCGGAAGAAAGTTCGAAGGAGCCTGTGTTCTTAAATAATCACATGCGGACACATAATGGC 811
Qy 541 AAATCGGGGGCCAGAAAGAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 600
Db 812 AAATCGGGGGCCAGAAAGAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 871
Qy 601 GAGGTGTCGAGGTGCAACGGGCGAGAGCATCTCCTCTCTTCAAAATCTGCAATGTTT 660
Db 872 GAGGTGTCGAGGTGCAACGGGCGAGAGCATCTCCTCTCTTCAAAATCTGCAATGTTT 931
Qy 661 TGTGCTTCTTATTTCCAAATTAAGAAAGTCTTAATTTAGAGCCGCAAGGTGCACACCAA 720
Db 932 TGTGCTTCTTATTTCCAAATTAAGAAAGTCTTAATTTAGAGCCGCAAGGTGCACACCAA 991
Qy 721 AAAACTGCTTTTCGTTACAGAGCGCGCAGACAGACTCTCCACAGGAGGAATGCGGTCC 780
Db 992 AAAACTGCTTTTCGTTACAGAGCGCGCAGACAGACTCTCCACAGGAGGAATGCGGTCC 1051
Qy 781 TCGAGGGAGGACTTCTCTGAGTTGTTCAAATTGAGACCAAAATCTCAACCTGAAACGGGG 840
Db 1052 TCGAGGGAGGACTTCTCTGAGTTGTTCAAATTGAGACCAAAATCTCAACCTGAAACGGGG 1111
Qy 841 AAGAGCCTGTGAGATGTCATCCCTCAGCTCGATCCGTTTCAACCTTCCAGGCTTGGCAG 900
Db 1112 AAGAGCCTGTGAGATGTCATCCCTCAGCTCGATCCGTTTCAACCTTCCAGGCTTGGCAG 1171
Qy 901 CTGGCTTACAAAGGAAAAAGTTGCCATTTGCCAAAGAAAGTGAAGGAATCGGGGCAAGAGGG 960
Db 1172 CTGGCTTACAAAGGAAAAAGTTGCCATTTGCCAAAGAAAGTGAAGGAATCGGGGCAAGAGGG 1231
Qy 961 AGCACCGCAACGACGATTCGAGTTCCGAGAAAGGAGCTTGGAGAAACAAAATAAGGGCAGT 1020
Db 1232 AGCACCGCAACGACGATTCGAGTTCCGAGAAAGGAGCTTGGAGAAACAAAATAAGGGCAGT 1291
Qy 1021 TGTGAGGCTCTCGCAAGAGAAAGAGTGAAGTCAACACTCCACGGCGAAGCGCCCTCC 1080
Db 1292 TGTGAGGCTCTCGCAAGAGAAAGAGTGAAGTCAACACTCCACGGCGAAGCGCCCTCC 1351
Qy 1081 GTGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAAAGCCCACTCACTGTCTCCAGTGC 1140
Db 1352 GTGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAAAGCCCACTCACTGTCTCCAGTGC 1411
Qy 1141 GGCAAAGCTTTAGAACTTACCAACAGCTGTGTTGCACTCCAGGGTCCACNAGAGGAC 1200
|||||

Db 1412 GGCAAAAGCTTTTGAAACCTTACCACCGAGTGGTCTTTGCACCTCCAGGGTCCACAAGAGGAC 1471
Qy 1201 CGGAGGCGCGCGGAGTGGCCACCACCATGCTCTGTGACGGGAGGCGAGCCGGGACCTGT 1260
Db 1472 CGGAGGCGCGCGGAGTGGCCACCACCATGCTCTGTGACGGGAGGCGAGCCGGGACCTGT 1531
Qy 1261 TCTCCTCACTCGCGCCCTCTGGATGAAATGAGGCCGTGGATCGAGGGGAAGGTGGT 1320
Db 1532 TCTCCTCACTCGCGCCCTCTGGATGAAATGAGGCCGTGGATCGAGGGGAAGGTGGT 1591
Qy 1321 TCTGAAGACGATCTGAGGATGGCTTCCGAGGAATCCATCTGGAATAAAAATGATGAT 1380
Db 1592 TCTGAAGACGATCTGAGGATGGCTTCCGAGGAATCCATCTGGAATAAAAATGATGAT 1651
Qy 1381 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGTGGAAGTTTTTC 1440
Db 1652 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGTGGAAGTTTTTC 1711
Qy 1441 CGTTCAAAATTTATTAACCTCAATATTTCACTCAGAACGATACAGGTGAAAAACCATACAAA 1500
Db 1712 CGTTCAAAATTTATTAACCTCAATATTTCACTCAGAACGATACAGGTGAAAAACCATACAAA 1771
Qy 1501 TGTGAATTTTGTGAATATGTCGAGCCGAGAGACATCTCTGAGGTATCACTTGGAGAGA 1560
Db 1772 TGTGAATTTTGTGAATATGTCGAGCCGAGAGACATCTCTGAGGTATCACTTGGAGAGA 1831
Qy 1561 CATCAAGGAAAAACAAACCGATGTTGCTGCTCAAGTCAAGAACGATGGTAAAAATCAG 1620
Db 1832 CATCAAGGAAAAACAAACCGATGTTGCTGCTGCTGAGTCAAGAACGATGGTAAAAATCAG 1891
Qy 1621 GACACTGAAGATGCACTATTAAACCGCTGACAGTGGCAGAACCAAAATTTGAAAGATTT 1680
Db 1892 GACACTGAAGATGCACTATTAAACCGCTGACAGTGGCAGAACCAAAATTTGAAAGATTT 1951
Qy 1681 TTTGATGTCGCAAGAGTGTACAGGAGTCCACTGCAAGAGAGCTTAAAGGATGCT 1740
Db 1952 TTTGATGTCGCAAGAGTGTACAGGAGTCCACTGCAAGAGAGCTTAAAGGATGCT 2011
Qy 1741 TCTGTTTTCAAGATGTTCTGGGAGGCTGCTCTCTCACCAGCACACAAAGATACTCAG 1800
Db 2012 TCTGTTTTCAAGATGTTCTGGGAGGCTGCTCTCTCACCAGCACACAAAGATACTCAG 2071
Qy 1801 GATTTCCATAAAAAATGCAAGTGTGATGAAAGTGAATAAAAACCTTACCCCT 1860
Db 2072 GATTTCCATAAAAAATGCAAGTGTGATGAAAGTGAATAAAAACCTTACCCCT 2131
Qy 1861 GCTTACCTGAGCTGTTAAAAAAGAGATCAGAGTTGAACTCAGGCAAAATAA CTTATC 1920
Db 2132 GCTTACCTGAGCTGTTAAAAAAGAGATCAGAGTTGAACTCAGGCAAAATAA CTTATC 2191
Qy 1921 TGTAGAACCAAGGCGGATGTTTACTCTCTCCGATGGCAGTACCAACCAAACTTCAA 1980
Db 2192 TGTAGAACCAAGGCGGATGTTTACTCTCTCCGATGGCAGTACCAACCAAACTTCAA 2251
Qy 1981 GTTAGCCCCCAAGAGAGCAAAACGAGACCGCAGCTGATCGAGATACAGGCCAAGTGTG 2040
Db 2252 GTTAGCCCCCAAGAGAGCAAAACGAGACCGCAGCTGATCGAGATACAGGCCAAGTGTG 2311
Qy 2041 GATTTGTCAGGAAAAACCTTTAAATTTATTCGTTGGGGCTCTTCAAAATGTCGCGCAAT 2100
Db 2312 GATTTGTCAGGAAAAACCTTTAAATTTATTCGTTGGGGCTCTTCAAAATGTCGCGCAAT 2371
Qy 2101 TCTTTGAGTAAAGTTTGAATTTCAAGTATCACTGTCTCATTTTGTACCTTCAAGACATTT 2160
Db 2372 TCTTTGAGTAAAGTTTGAATTTCAAGTATCACTGTCCATTTTGTACCTTCAAGACATTT 2431
Qy 2161 TATCCAGAGTTTAAATGATGCAACGAGACTGGAGCATAAATACAATCTCTGACGTTCA 2220
Db 2432 TATCCAGAGTTTAAATGATGCAACGAGACTGGAGCATAAATACAATCTCTGACGTTCA 2491
Qy 2221 AAAAAGCTGTCAAAACAGTCTCTCTTGAAGTGCAGTACCGGATGCCCGGACGTTG 2280
Db 2492 AAAAAGCTGTCAAAACAGTCTCTCTTGAAGTGCAGTACCGGATGCCCGGACGTTG 2551

| | | | |
|----|----------------------------|--|------|
| XX | Sequence | 5632 BP; 1637 A; 1167 C; 1257 G; 1571 T; 0 U; 0 Other; | |
| SQ | Query Match | 95.4%; Score 3040.6; DB 13; Length 5632; | |
| | Best Local Similarity | 95.8%; Pred. No. 0; | |
| | Matches 3180; Conservative | 6; Mismatches 0; Indels 133; Gaps 1; | |
| Qy | 1 | ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAAATGATCATGGATGGG | 60 |
| Db | 272 | ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAAATGATCATGGATGGG | 331 |
| Qy | 61 | CCAAAGTGATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG | 120 |
| Db | 332 | CCAAAGTGATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG | 391 |
| Qy | 121 | AAAGGGACCGCTGTGTTCATTCGGAGCTACACAGAAAATAATGTCATCCAAATCGAG | 180 |
| Db | 392 | AAAGGGACCGCTGTGTTCATTCGGAGCTACACAGAAAATAATGTCATCCAAATCGAG | 451 |
| Qy | 181 | GGGTATATGCCCTTGGATGTGATGTTCTGCAGCCAGACCTTTCACACATTCAGAAAGACTT | 240 |
| Db | 452 | GGGTATATGCCCTTGGATGTGATGTTCTGCAGCCAGACCTTTCACACATTCAGAAAGACTT | 511 |
| Qy | 241 | AATAAACAATGCTTAAATGCAACACCGGCTTACCCTCTGTGAACACGACAGTTCTTCCGGTT | 300 |
| Db | 512 | AATAAACAATGCTTAAATGCAACACCGGCTTACCCTCTGTGAACACGACAGTTCTTCCGGTT | 571 |
| Qy | 301 | GAACAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGCAGACAGACCTCCCAAGAA | 360 |
| Db | 572 | GAACAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGCAGACAGACCTCCCAAGAA | 631 |
| Qy | 361 | AAGAAATTCAGAGAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTGGCT | 420 |
| Db | 632 | AAGAAATTCAGAGAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTGGCT | 691 |
| Qy | 421 | TTTGATGTGAGATCCACATGAGAAACACAAAAGATTTCTTCACTTACGGGTGAAACATG | 480 |
| Db | 692 | TTTGATGTGAGATCCACATGAGAAACACAAAAGATTTCTTCACTTACGGGTGAAACATG | 751 |
| Qy | 481 | TGCGGAAGAGMTTSRRSSAGCCTGGTTTCTTAAATATCATGCGGACACATTAATGGC | 540 |
| Db | 752 | TGCGGAAGAGMTTCAAGAGCCTGGTTTCTTAAATATCATGCGGACACATTAATGGC | 811 |
| Qy | 541 | AAATCGGGGGCCAGAAAGAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC | 600 |
| Db | 812 | AAATCGGGGGCCAGAAAGAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC | 871 |
| Qy | 601 | GAGTGTGTCAGGTGCACGGCGCGAGAGATCTCCTCTCTTACAAATCTGCATGGTT | 660 |
| Db | 872 | GAGTGTGTCAGGTGCACGGCGCGAGAGATCTCCTCTCTTACAAATCTGCATGGTT | 931 |
| Qy | 661 | TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA | 720 |
| Db | 932 | TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA | 991 |
| Qy | 721 | AAACTGCTTTCCGTACACAGCAGCGCAGACAGACTCTCCACAGGAGGAAATGCCGTCC | 780 |
| Db | 992 | AAACTGCTTTCCGTACACAGCAGCGCAGACAGACTCTCCACAGGAGGAAATGCCGTCC | 1051 |
| Qy | 781 | TCGAGGGAGGACTTCCCTCAGTGTGTTCAACTTGGAGACCAAAATCTCACCTGAAACGGGG | 840 |
| Db | 1052 | TCGAGGGAGGACTTCCCTCAGTGTGTTCAACTTGGAGACCAAAATCTCACCTGAAACGGGG | 1111 |
| Qy | 841 | AAGAGCCTGTCCAGATGATCCCTCAGCTCCGATCCGTTCCACCTTCCAGGCTTGGCAG | 900 |
| Db | 1112 | AAGAGCCTGTCCAGATGATCCCTCAGCTCCGATCCGTTCCACCTTCCAGGCTTGGCAG | 1171 |
| Qy | 901 | CTGGCTACCAAGGAAAGTTGCCATTTGCCAATGGAAGTGAAGAAATCGGGCAAGAGGG | 960 |
| Db | 1172 | CTGGCTACCAAGGAAAGTTGCCATTTGCCAATGGAAGTGAAGAAATCGGGCAAGAGGG | 1231 |
| Qy | 961 | AGCACCGACACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATAAGGGCAGT | 1020 |
| Db | | | |

| | | | |
|----|------|--|------|
| Db | 1232 | AGCACCGACAAACGACGATTCGAGTTCGAGAAAGAGCTTGGAGAGAAACAAATAAGGGCAGT | 1291 |
| Qy | 1021 | TGTGCAGGCTCTCGCAAGAGAAAGAGAGTGCACAAACACTCCACGCGGAGAGCCCTCC | 1080 |
| Db | 1292 | TGTGCAGGCTCTCGCAAGAGAAAGAGAGTGCACAAACACTCCACGCGGAGAGCCCTCC | 1351 |
| Qy | 1081 | GTGACGCGGATCCAAAGTTTACCCAGTACGAGGAGAGCCCACTCACTGCTCCGAGTGC | 1140 |
| Db | 1352 | GTGACGCGGATCCAAAGTTTACCCAGTACGAGGAGAGCCCACTCACTGCTCCGAGTGC | 1411 |
| Qy | 1141 | GGCAAGCTTTTCAGAACTTACCAACAGCTGGTCTTGCACTCCAGGGTCCACAGAAAGAC | 1200 |
| Db | 1412 | GGCAAGCTTTTCAGAACTTACCAACAGCTGGTCTTGCACTCCAGGGTCCACAGAAAGAC | 1471 |
| Qy | 1201 | CGAGGGCCGCGCGGAGTCCGCCACCATGTCTGTGACCGGAGGAGCGGGGACGTGT | 1260 |
| Db | 1472 | CGAGGGCCGCGCGGAGTCCGCCACCATGTCTGTGACCGGAGGAGCGGGGACGTGT | 1531 |
| Qy | 1261 | TCTCTGACCTCGCCGCCCTCTGGATGAAATGGAGCCGTGGATCGAGGGGAAAGGTGGT | 1320 |
| Db | 1532 | TCTCTGACCTCGCCGCCCTCTGGATGAAATGGAGCCGTGGATCGAGGGGAAAGGTGGT | 1591 |
| Qy | 1321 | TCTGAACAGCGATCTGAGGATGGCTTCCCGAAGGAATCCATCTGGATATAAATATGATGAT | 1380 |
| Db | 1592 | TCTGAACAGCGATCTGAGGATGGCTTCCCGAAGGAATCCATCTGGATATAAATATGATGAT | 1651 |
| Qy | 1381 | GGAGAAAATAAACAATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAGTTTTTC | 1440 |
| Db | 1652 | GGAGAAAATAAACAATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAGTTTTTC | 1711 |
| Qy | 1441 | CGTTCAAATTTATACCTCAATATTTCTCAGAACCGATACAGGTGAAAAACCATACAAA | 1500 |
| Db | 1712 | CGTTCAAATTTATACCTCAATATTTCTCAGAACCGATACAGGTGAAAAACCATACAAA | 1771 |
| Qy | 1501 | TGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCTGAGGTATCACTTCGAGAGA | 1560 |
| Db | 1772 | TGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCTGAGGTATCACTTCGAGAGA | 1831 |
| Qy | 1561 | CATCAAGGAAAAACAAACCGATTTGTCTGCTCAAGTCAAGAACGATGGTAAAAATCAG | 1620 |
| Db | 1832 | CATCAAGGAAAAACAAACCGATTTGTCTGCTGCTGCTCAAGAACGATGGTAAAAATCAG | 1891 |
| Qy | 1621 | GACACTGAAGATGCACTATTAAACCGCTGACAGTGGCNAACCAAAATTTGAAAAGATTT | 1680 |
| Db | 1892 | GACACTGAAGATGCACTATTAAACCGCTGACAGTGGCNAACCAAAATTTGAAAAGATTT | 1951 |
| Qy | 1681 | TTTGATGTGTCGAAGATGTTACAGGAGTCCACCTGCAGAGAGCTTAAAGAGATGCT | 1740 |
| Db | 1952 | TTTGATGTGTCGAAGATGTTACAGGAGTCCACCTGCAGAGAGCTTAAAGAGATGCT | 2011 |
| Qy | 1741 | TCTGTTTTTCAGAAATGTTCTGGGCGAGCTGCTCTCTCACACGACACAAAGATCTCAG | 1800 |
| Db | 2012 | TCTGTTTTTCAGAAATGTTCTGGGCGAGCTGCTCTCTCACACGACACAAAGATCTCAG | 2071 |
| Qy | 1801 | GATTTCCATAAAAAATGCAGCTGATGACAGTCTGATTAAGTGAATAAAAAACCTTACCCCT | 1860 |
| Db | 2072 | GATTTCCATAAAAAATGCAGCTGATGACAGTCTGATTAAGTGAATAAAAAACCTTACCCCT | 2131 |
| Qy | 1861 | GCTTACCTGGACCTGTTAAAAAGAGATCAGCAGTTGAAAACCTCAGGCAAAATAACCTCATC | 1920 |
| Db | 2132 | GCTTACCTGGACCTGTTAAAAAGAGATCAGCAGTTGAAAACCTCAGGCAAAATAACCTCATC | 2191 |
| Qy | 1921 | TGTAGAACCAAGGGCGGATGTTTACTCTCTCCGATGGCAGTACACCCATTAACCTTGAA | 1980 |
| Db | 2192 | TGTAGAACCAAGGGCGGATGTTTACTCTCTCCGATGGCAGTACACCCATTAACCTTGAA | 2251 |
| Qy | 1981 | GTTAGCCCAAGAGAGAGCAACCGAGACCGAGCTGATCTGCAGATACAGGCCAAGTGTG | 2040 |
| Db | 2252 | GTTAGCCCAAGAGAGAGCAACCGAGACCGAGCTGATCTGCAGATACAGGCCAAGTGTG | 2311 |
| Qy | 2041 | GATTTGTACGAAACCTTTTAAATTTATCCGTGGGGCTCTTCAAAATTTCCCGCAATT | 2100 |
| Db | 2312 | GATTTGTACGAAACCTTTTAAATTTATCCGTGGGGCTCTTCAAAATTTCCCGCAATT | 2371 |

QY 2101 TCCTTTGAGTAAAGTTTGGATTCCTCAAGTATCACCTGTCATTTTGTACCTTCAAGACATTT 2160
DB |||||
QY 2372 TCCTTTGAGTAAAGTTTGGATTCCTCAAGTATCACCTGTCATTTTGTACCTTCAAGACATTT 2431
DB |||||
QY 2161 TATCAGAAAGTTTAAATGATGACACAGAGACTGGAGCATATAAATAACAATCTTGACCTTCAT 2220
DB |||||
QY 2432 TATCAGAAAGTTTAAATGATGACACAGAGACTGGAGCATATAAATAACAATCTTGACCTTCAT 2491
DB |||||
QY 2221 AAAAATGTGCGAAACAAAGTCTTGTCTAGAGTCAAGTACCGGATCGCCGACAGGTTG 2280
DB |||||
QY 2492 AAAAATGTGCGAAACAAAGTCTTGTCTAGAGTCAAGTACCGGATCGCCGACAGGTTG 2551
DB |||||
QY 2281 CTGGGAAAGATGTGCTCCCTCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTCTCTTC 2340
DB |||||
QY 2552 CTGGGAAAGATGTGCTCCCTCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTCTCTTC 2611
DB |||||
QY 2341 CGCGGCGAGTCCAAATCCCTGCGCATCTGCGAAGGGGAAAGCAGAGCCCTCTGGGCGAGGC 2400
DB |||||
QY 2612 CGCGGCGAGTCCAAATCCCTGCGCATCTGCGAAGGGGAAAGCAGAGCCCTCTGGGCGAGGC 2671
DB |||||
QY 2401 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCCAAGTAACCTGAAGTCC 2460
DB |||||
QY 2672 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCCAAGTAACCTGAAGTCC 2731
DB |||||
QY 2461 CACAGCACACAGCAATGTGGGGTCCAAAGGGCGCCACAGGCAACGCAATCTGAG 2520
DB |||||
QY 2732 CACAGCACACAGCAATGTGGGGTCCAAAGGGCGCCACAGGCAACGCAATCTGAG 2791
DB |||||
QY 2521 ATGTTTCTTAAACCAAGTGTTCCTCTGACCGGATAGACAAAGAACCCGAGACAAA 2580
DB |||||
QY 2792 ATGTTTCTTAAACCAAGTGTTCCTCTGACCGGATAGACAAAGAACCCGAGACAAA 2851
DB |||||
QY 2581 TTGAACCTCTTCCAGTAGTCTCTCTGAGCCACCTCGGAGAGTAAACATCAATGCT 2640
DB |||||
QY 2852 TTGAACCTCTTCCAGTAGTCTCTCTGAGCCACCTCGGAGAGTAAACATCAATGCT 2911
DB |||||
QY 2641 TCCATCGATACCCCGCCAGAACACACAGCCGTCGGACCTCCGGGAGAGACTATTTC 2700
DB |||||
QY 2912 TCCATCGATACCCCGCCAGAACACACAGCCGTCGGACCTCCGGGAGAGACTATTTC 2971
DB |||||
QY 2701 TGTATCGAGTCCAGCAATCTGACGAGAAATTTGTTGAGCCCTCTCCAAAGAGCTG 2760
DB |||||
QY 2972 TGTATCGAGTCCAGCAATCTGACGAGAAATTTGTTGAGCCCTCTCCAAAGAGCTG 3031
DB |||||
QY 2761 AAGTCCAGCGTGTGCTGCTGACGTTGACAGCCCGGGCCAAATACAGAGAGGCTAT 2820
DB |||||
QY 3032 AAGTCCAGCGTGTGCTGCTGACGTTGACAGCCCGGGCCAAATACAGAGAGGCTAT 3091
DB |||||
QY 2821 GACCTTCCAGTACATATGTCAGAGGCATCATCATCTGTTTACCGAGGACTGTGTG 2880
DB |||||
QY 3092 GACCTTCCAGTACATATGTCAGAGGCATCATCATCTGTTTACCGAGGACTGTGTG 3151
DB |||||
QY 2881 TATCGTCCAGCGCTGCTGCCAAACCAAGGTTCTTGAGCTCCAGCGAGTCTGATTC 2940
DB |||||
QY 3152 TATCGTCCAGCGCTGCTGCCAAACCAAGGTTCTTGAGCTCCAGCGAGTCTGATTC 3211
DB |||||
QY 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGGCTCCGGGCCACTTTACACTTGTGTG 3000
DB |||||
QY 3212 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGGCTCCGGGCCACTTTACACTTGTGTG 3271
DB |||||
QY 3001 CTGCTGTGATGTCAGATTCAGTCTGAGCTGAGTGTAG----- 3034
DB |||||
QY 3272 CTGCTGTGATGTCAGATTCAGTCTGAGCTGAGTGTAG----- 3331
DB |||||
QY 3035 ----- 3034
DB |||||
QY 3332 CACTTATCTAACAGCATGCGCAAAAGAGAACTATGAGAAATTTATTGGGAATGCAT 3391
DB |||||
QY 3035 ----- 3034
DB |||||
QY 3392 TATGACCAATGACAAAAAACTTGATTCATTAATTAGGGGAAAAAGGCTCTTGGTGG 3451
DB |||||

QY 3048 ATGTCAGTGTCTTACTCCCATGAAATTAATTTTACTTCATCCTTTGAGAGCGAATGCT 3107
DB |||||
QY 3452 ATGTCAGTGTCTTACTCCCATGAAATTAATTTTACTTCATCCTTTGAGAGCGAATGCT 3511
DB |||||
QY 3108 GAAAGCTACTGAAATAAGCTGTGATTTGACTGTACATAAAACATATGAGGAATCTGCAAG 3167
DB |||||
QY 3512 GAAAGCTACTGAAATAAGCTGTGATTTGACTGTACATAAAACATATGAGGAATCTGCAAG 3571
DB |||||
QY 3168 GAAACTACATGTTGTGTAA 3186
DB |||||
QY 3572 GAACTACATGTTGTGTAA 3590
DB |||||
RESULT 6
ADR65858
ID ADR65858 standard; DNA; 5632 BP.
XX
AC ADR65858;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human prostatic carcinoma derived DNA SEQ ID 54 #1.
XX
KW human; cytostatic; diagnosis; prostatic cancer;
KW differential expression analysis; ds.
XX
OS Homo sapiens.
XX
PN WO2004076614-A2.
XX
PD 10-SEP-2004.
XX
PF 22-FEB-2004; 2004WO-DE000433.
XX
PR 27-FEB-2003; 2003DE-01009985.
PR 14-MAY-2003; 2003DE-01022134.
XX
PA (HINZ/) HINZMANN B.
PA (DAHL/) DAHL E.
PA (ROSE/) ROSENTHAL A.
PA (HERM/) HERMANN K.
PA (PILA/) PILARSKY C.
XX
PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarczyk C, Specht T;
PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
PI Xinzhong L, Staub E;
XX
DR WPI; 2004-653386/63.
XX
PT New nucleic acids, and encoded proteins, from prostatic cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
agents.
XX
PS Claim 1; Page 253-254; 1607pp; German.
XX
CC This invention describes novel cytostatic polynucleotide and polypeptide
CC sequences which can be used in a method for diagnosing prostatic cancer
CC or the risk of developing prostatic cancer. Diagnosis is based on
CC determining over transcription or over expression of the sequences in
CC prostatic tissue. Screening for inhibitors of the sequences or detection
CC substances involves a binding assay, any compounds that bind are
CC selected, optionally after deconvolution of mixtures. Detection of a
CC predetermined minimum level of the reporter indicates the presence of
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
CC short-interfering RNA or ribozymes; an organic molecule of molecular
CC weight below 5000, preferably 300, that binds to the polypeptide; an
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human
CC (monoclonal) antibody directed against Ab or any of the above derivatised
CC with a reporter group, cell toxin, immunostimulatory molecules and/or
CC radioisotope. The polynucleotides are identified in human prostatic
CC cancer by differential expression analysis, using DNA microarrays,
CC between normal and tumorous tissues, with (over)expression being detected

CC by quantitative PCR. Analysis of prostatic cancer samples showed that
CC CD24 was upregulated in many of them. Sections of tissue, isolated from
CC prostatic cancer patients, or subjects at risk, were incubated
CC sequentially with anti-human CD4 murine monoclonal antibodies;
CC biotinylated second antibody; streptavidin-conjugated horseradish
CC peroxidase and then diaminobenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. ADR65805-ADR6954 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.
XX
SQ Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;
Query Match 95.4%; Score 3040.6; DB 13; Length 5632;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;
Qy 1 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGG 60
Db 272 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGG 331
Qy 61 CCAGAAGTGATGGCAGCTCTCTTGGCAGTCCGATCGATGGAGATGCCCTTGTCAATG 120
Db 332 CCAGAAGTGATGGCAGCTCTCTTGGCAGTCCGATCGATGGAGATGCCCTTGTCAATG 391
Qy 121 AAAGGGACCGCTGTGTTCCATTCGGAGCTACACAAGAAAAAATGTCCATCCAAATCGAG 180
Db 392 AAAGGGACCGCTGTGTTCCATTCGGAGCTACACAAGAAAAAATGTCCATCCAAATCGAG 451
Qy 181 GGGTATATGCCCTTGGATGTGATGTTCTGACGCGACACTTTCACACATTCAGAAAGACCTT 240
Db 452 GGGTATATGCCCTTGGATGTGATGTTCTGACGCGACACTTTCACACATTCAGAAAGACCTT 511
Qy 241 AATAAATGCTCTTAATGCAACACCGGCTTACCCTCTGTGTGAACGACGAGTTCTTCGGGTT 300
Db 512 AATAAATGCTCTTAATGCAACACCGGCTTACCCTCTGTGTGAACGACGAGTTCTTCGGGTT 571
Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 360
Db 572 GAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 631
Qy 361 AAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGACAGACATTTAGAGTCGCT 420
Db 632 AAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGACAGACATTTAGAGTCGCT 691
Qy 421 TTTGATGTTGAGATCCATGAGAAACACAAAAGATTCTTTTCACTTACGGGTGTAAACATG 480
Db 692 TTTGATGTTGAGATCCATGAGAAACACAAAAGATTCTTTTCACTTACGGGTGTAAACATG 751
Qy 481 TGCGAAGCAAGMTTSRRSSAGCCTTGGTTTCTTAAATAATCATGCGGACACATAATGGC 540
Db 752 TGCGAAGCAAGATTCAAGAGGCTTGGTTTCTTAAATAATCATGCGGACACATAATGGC 811
Qy 541 AAATCGGGGGCCAGAAAGAACTGCAGCAAGGCTTTGGAGAGTAGTCCAGCAACGATCAAC 600
Db 812 AAATCGGGGGCCAGAAAGAACTGCAGCAAGGCTTTGGAGAGTAGTCCAGCAACGATCAAC 871
Qy 601 GAGTGTGTCAGGTGCACGGGCCGAGAGCATCTCTCTCTTCAATAATCTGCATGGTT 660
Db 872 GAGTGTGTCAGGTGCACGGGCCGAGAGCATCTCTCTCTTCAATAATCTGCATGGTT 931
Qy 661 TGTGGCTTCTTATTTTCCAAATAAGAAAGTCTAATTGAGCAGCGGAGGTGCACACCAA 720
Db 932 TGTGGCTTCTTATTTTCCAAATAAGAAAGTCTAATTGAGCAGCGGAGGTGCACACCAA 991
Qy 721 AAAAATGCTTTTCGGTACCAGCAGCGGACAGACAGACTCTCCACAGGAGGAATGCCGTCC 780
Db 992 AAAAATGCTTTTCGGTACCAGCAGCGGACAGACAGACTCTCCACAGGAGGAATGCCGTCC 1051
Qy 781 TCGAGGAGGAGTCTCTCGAGTTGTTCAATTGAGACCAAAATCTCACCTCGAAACGGGG 840
|||||

Db 1052 TCGAGGGAGGACTTCTCGCAGTTGTTTCAACTTGTGAGACCAAAATCTCACCTCGAAACGGGG 1111
Qy 841 AAGAAGCCTGTGATGATCATCCCTCAGCTCGATCCGTTCCACCACTTCCAGGCTTGGCAG 900
Db 1112 AAGAAGCCTGTGATGATCATCCCTCAGCTCGATCCGTTCCACCACTTCCAGGCTTGGCAG 1171
Qy 901 CTGGCTACCAAGAAAGTTGCCATTTGGCCAAAGATGAAGGAATCGGGCAAGAAGGG 960
Db 1172 CTGGCTACCAAGAAAGTTGCCATTTGGCCAAAGATGAAGGAATCGGGCAAGAAGGG 1231
Qy 961 AGCACCCACACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGCGAGT 1020
Db 1232 AGCACCCACACGACGATTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGCGAGT 1291
Qy 1021 TGTGAGCGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCACGCGCAAGCCCTCC 1080
Db 1292 TGTGAGCGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCACGCGCAAGCCCTCC 1351
Qy 1081 GTGAGCGGGATCCCAAGTTTACCAGTAGCAAGAGAGAGGCCACTCACTGCTCCGAGTGC 1140
Db 1352 GTGAGCGGGATCCCAAGTTTACCAGTAGCAAGAGAGAGGCCACTCACTGCTCCGAGTGC 1411
Qy 1141 GGCAAGCCTTTCAGAACCTTACCACAGCTGTCTTGCACTCCAGGCTCCACAAGAGGAC 1200
Db 1412 GGCAAGCCTTTCAGAACCTTACCACAGCTGTCTTGCACTCCAGGCTCCACAAGAGGAC 1471
Qy 1201 CGAGGGCGGGCGGAGTCCGCCACCATGTCTGTGACGCGGAGCGACCCGGGACCTGT 1260
Db 1472 CGAGGGCGGGCGGAGTCCGCCACCATGTCTGTGACGCGGAGCGACCCGGGACCTGT 1531
Qy 1261 TCTCCTGACCTTCGCGGCCCTCTCGGATGAAATGGAGCGGTGGATCGAGGGGAAGTGGT 1320
Db 1532 TCTCCTGACCTTCGCGGCCCTCTCGGATGAAATGGAGCGGTGGATCGAGGGGAAGTGGT 1591
Qy 1321 TCTCAAGACCGATCTGAGGATGGCTTCCGAGGAAATCCATCTGGATATAAATGATGAT 1380
Db 1592 TCTCAAGACCGATCTGAGGATGGCTTCCGAGGAAATCCATCTGGATATAAATGATGAT 1651
Qy 1381 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGGAAAGTTTTTC 1440
Db 1652 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGGAAAGTTTTTC 1711
Qy 1441 CGTTCAAATTTATTAACCTCAATATTCATCTCAGAACGATACAGGTGAAAAACCATACAA 1500
Db 1712 CGTTCAAATTTATTAACCTCAATATTCATCTCAGAACGATACAGGTGAAAAACCATACAA 1771
Qy 1501 TGTCAATTTTGTGAATATGCTGACGCGCCAGAGACATCTCTGAGGTATCATCTTGAGAGA 1560
Db 1772 TGTGAATTTTGTGAATATGCTGACGCGCCAGAGACATCTCTGAGGTATCATCTTGAGAGA 1831
Qy 1561 CATCAAGGAAAAACAAACCGATGTTGCTGCTGAAAGTCAAGAACGATGGTAAAAATCAG 1620
Db 1832 CATCAAGGAAAAACAAACCGATGTTGCTGCTGAAAGTCAAGAACGATGGTAAAAATCAG 1891
Qy 1621 GACATGAAGATGACATTAATTAACCGCTGACAGTGGCGAAACCAAAAAATTTGAAAGATTT 1680
Db 1892 GACATGAAGATGACATTAATTAACCGCTGACAGTGGCGAAACCAAAAAATTTGAAAGATTT 1951
Qy 1681 TTTGATGTCGCAAGATGTTTACAGGCGATCCACTGCAAGAGAGCTTAAGAGAGATGCT 1740
Db 1952 TTTGATGTCGCAAGATGTTTACAGGCGATCCACTGCAAGAGAGCTTAAGAGAGATGCT 2011
Qy 1741 TCTGTTTTTTCAGAAATGTTCTGGGCGAGCTGCTCTCTCACGACACACAAAGATCTCAG 1800
Db 2012 TCTGTTTTTTCAGAAATGTTCTGGGCGAGCTGCTCTCTCACGACACACAAAGATCTCAG 2071
Qy 1801 GATTTCCATAAAAAATGCGAGTGTGAAGTCTGATTAAGTGAATAAAAAATTTACCCCT 1860
Db 2072 GATTTCCATAAAAAATGCGAGTGTGACAGTCTGATTAAGTGAATAAAAAATTTACCCCT 2131
Qy 1861 GCTTACCTGGACCTGTTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAAATCTCATC 1920
Db 2132 GCTTACCTGGACCTGTTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAAATCTCATC 2191
|||||

1921 TGTAAGCAACAGCGGATGTTACTCCTCCGATGGCAGTACCAACCAATACCTTCAA 1980
1922 TGTAAGCAACAGCGGATGTTACTCCTCCGATGGCAGTACCAACCAATACCTTCAA 2251
1981 GTTAGCCCCAAGAGAGCAACAGGAGACCGCAGCTGACTGCAGATACAGGCCAAGTGTG 2040
2252 GTTAGCCCCAAGAGAGCAACAGGAGACCGCAGCTGACTGCAGATACAGGCCAAGTGTG 2311
2041 GATTGTACAGAAACCTTTAAATTTATCCGTGGGGCTCTTCAATGTCGCCGCAATT 2100
2312 GATTGTACAGAAACCTTTAAATTTATCCGTGGGGCTCTTCAATGTCGCCGCAATT 2371
2101 TCTTTGAGTAAAGTTGATTCCTCAAGTATCACTGTCCTTGTGATCTTCAAGACATTT 2160
2372 TCTTTGAGTAAAGTTGATTCCTCAAGTATCACTGTCCTTGTGATCTTCAAGACATTT 2431
2161 TATCCAGAAGTTTAAATGATGACACAGAGACTGGAGCATAAATACATCTTCAAGTTCAT 2220
2432 TATCCAGAAGTTTAAATGATGACACAGAGACTGGAGCATAAATACATCTTCAAGTTCAT 2491
2221 AAAAATCTGTGAAACAGTCTTGTGATAGAGTCGACGTACCGGATGCCCGCAGCGTTG 2280
2492 AAAAATCTGTGAAACAGTCTTGTGATAGAGTCGACGTACCGGATGCCCGCAGCGTTG 2551
2281 CTGGGAAGATGTCCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTC 2340
2552 CTGGGAAGATGTCCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTC 2611
2341 CCGGCGCAGTCCAAATCCCTGCAATCTGCAAGGAGGAGAGAGCCCTCTGGGCGAGGC 2400
2612 CCGGCGCAGTCCAAATCCCTGCAATCTGCAAGGAGGAGAGAGCCCTCTGGGCGAGGC 2671
2401 AAGGCCCTCTGACTTCAGGAGTACACTCTAGCACTTTAGCCCCCAAGTAACTGAAGTCC 2460
2672 AAGGCCCTCTGACTTCAGGAGTACACTCTAGCACTTTAGCCCCCAAGTAACTGAAGTCC 2731
2461 CACAGCACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACAGGCAACAGCAATCTGAG 2520
2732 CACAGCACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACAGGCAACAGCAATCTGAG 2791
2521 ATGTTTCTTAAACCAAGTGTTCCTCTGCAAGGAGTAAAGCAAAAGACCCGAGCAAAA 2580
2792 ATGTTTCTTAAACCAAGTGTTCCTCTGCAAGGAGTAAAGCAAAAGACCCGAGCAAAA 2851
2581 TTGAAACCTCTTCCAGTAGCTCTTCTCAGCCCCACCTCTCGCAGCAGTAAATCAATGTT 2640
2852 TTGAAACCTCTTCCAGTAGCTCTTCTCAGCCCCACCTCTCGCAGCAGTAAATCAATGTT 2911
2641 TCCATCGACTACCCCGCAAGAACAGACAGCCGCTGGGCACTCTCGGGAAGAGACTATTTTC 2700
2912 TCCATCGACTACCCCGCAAGAACAGACAGCCGCTGGGCACTCTCGGGAAGAGACTATTTTC 2971
2701 TGTATCGAGTGCCAGCAATACTGACAGCAAAATTTGTGAGCCCTTCCAAAAGACTG 2760
2972 TGTATCGAGTGCCAGCAATACTGACAGCAAAATTTGTGAGCCCTTCCAAAAGACTG 3031
2761 AAGTCCAGCGTGTGTCCTTTCAGCTTGAACAGCCCGGGGCAATTAACAGAGAGGCTAT 2820
3032 AAGTCCAGCGTGTGTCCTTTCAGCTTGAACAGCCCGGGGCAATTAACAGAGAGGCTAT 3091
2821 GACCTTCCAGTACCATATGTTAGAGGATCAATCATCTGTTTACCGAGAGACTGTGTG 2880
3092 GACCTTCCAGTACCATATGTTAGAGGATCAATCATCTGTTTACCGAGAGACTGTGTG 3151
2881 TATCCGTGCGAGGCGCTCCCTCCCAAAACCAAGTTCTGAGTCCAGCGAGTCCGATTTCT 2940
3152 TATCCGTGCGAGGCGCTCCCTCCCAAAACCAAGTTCTGAGTCCAGCGAGTCCGATTTCT 3211
2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGCTCCGGGCCACTTTACACTTGTGTG 3000
3212 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGCTCCGGGCCACTTTACACTTGTGTG 3271

3001 CCTGCTGCTAGTCCAGCATCCAGCTCCGACGTTAG----- 3034
3272 CCTGCTGCTAGTCCAGCATCCAGCTCCGACGTTAGAAAGGAAAGCCCTGTGTATATCAA 3331
3035 ----- 3034
3332 CACTTATCTAACAGCATGGCACAAAGAGAAACTATGAGAAATTTTATTTGGGAATCCACAT 3391
3035 -----AAGTCTTTGGTGG 3047
3392 TATCGACCAATGACAAAAAACTTGATTCAATATTAGGGGGGAAAAAAGGCTTTGGTGG 3451
3048 ATGTCAGTGTCTTACTTCCCATGAAATTAATTTTACTTTCATCTTTGAGAGCGGAATGCT 3107
3452 ATGTCAGTGTCTTACTTCCCATGAAATTAATTTTACTTTCATCTTTGAGAGCGGAATGCT 3511
3108 GAAAGCTACTGAAATAAGCTGTGATTGTACTGTACATAAAACATATGAGGAATCTGCAAG 3167
3512 GAAAGCTACTGAAATAAGCTGTGATTGTACTGTACATAAAACATATGAGGAATCTGCAAG 3571
3168 GAACACTACAGTTGTGTAA 3186
3572 GAACACTACAGTTGTGTAA 3590

RESULT 7

ADP07267
ID ADP07267 standard; DNA; 5653 BP.
XX
AC ADP07267;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human ZNF217 DNA.
XX
KW ds; proliferative disease; breast; methylation; CpG; bisulfite; human.
XX
OS Homo sapiens.
XX
PN DE10255104-A1.
XX
PD 11-MAR-2004.
XX
PF 26-NOV-2002; 2002DE-01055104.
XX
PR 27-AUG-2002; 2002DE-01039313.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Maier S;
XX
DR WPI; 2004-284340/27.
XX
DR GENBANK; NM_006526.
XX
PT Analyzing proliferative diseases of breast cells, useful e.g. for
PT diagnosis, prognosis and treatment of breast cancer, by determining
PT methylation status of specific genes.
XX
PS Claim 1; Page; 22pp; German.
XX
CC This invention describes a novel method of analysing proliferative
CC diseases of breast cells by determining the methylation status of certain
CC genes. The invention also describes nucleic acids or their complements,
CC oligomers, especially oligonucleotides or peptide nucleic acid oligomers,
CC that hybridise to, or are identical with, any of the nucleic acids, the
CC preparation of an oligomer array for analysing proliferative diseases of
CC breast cells that are associated with the methylation status of CpG
CC dinucleotides of the genes by bonding at least one oligomer to a solid
CC phase and a kit comprising a bisulfite reagent and the oligomer. The
CC nucleic acids are genomic sequences (5' - and/or regulatory and/or CpG-
CC rich regions). The base sequence of the oligomer includes at least one
CC CpG island, especially with C in the middle third of the sequence. The
CC process involves a genomic DNA sample treated chemically, specifically

CC with a bisulfite reagent, to convert non-methylated C to uracil or some
CC other base having base-pairing properties different from those of C.
CC Fragments of the treated DNA are then amplified, using primers and a
CC polymerase and the methylation status of the genomic CpG dinucleotides is
CC determined by analysis of the amplicons, particularly by hybridisation to
CC the oligomer, optionally with extension of the hybridised oligomer by at
CC least one base, or detection is by sequencing. The amplification may use
CC methylation-specific primers. Alternatively, genomic DNA is extracted
CC from a sample and digested with methylation-specific restriction enzymes,
CC then the digestion fragments detected, optionally after amplification. In
CC either method, more than 10 fragments of 100-200 bp are amplified in a
CC single reaction vessel, using a heat-resistant DNA polymerase in PCR. The
CC amplicons carry detectable markers, e.g. fluorophores, radioisotopes
CC and/or releasable fragments of known mass that can be detected by mass
CC spectrometry. The method is used for characterisation, classification,
CC differentiation, staging, phase-estimation, diagnosis and/or therapy of
CC proliferative diseases of breast cells. The method provides very specific
CC classification of proliferative diseases, allowing better treatment. It
CC can both characterise methylation status and detect single-nucleotide
CC polymorphisms. This sequence represents human gene used to illustrate the
CC method of the invention. NOTE: This sequence does not appear in the
CC printed specification but has been retrieved from Genbank.

xx
SQ Sequence 5653 BP; 1656 A; 1167 C; 1258 G; 1572 T; 0 U; 0 Other;

Query Match 95.4%; Score 3040.6; DB 12; Length 5653;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;

Qy 1 ATGCAATCGAAGTGACAGGAAACATGCCAACTCAATCCCTTTAATGTACATGATGGG 60
Db 272 ATGCAATCGAAGTGACAGGAAACATGCCAACTCAATCCCTTTAATGTACATGATGGG 331
Qy 61 CCAGAGTGATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGTGCCTTGTCAATG 120
Db 332 CCAGAGTGATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGTGCCTTGTCAATG 391
Qy 121 AAAGGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCCATCCAAATCGAG 180
Db 392 AAAGGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCCATCCAAATCGAG 451
Qy 181 GGGTATATGCCCTTGGATTGTCATGTTCTGACGCCAGACCTTTCACACATTCAGAGACCTT 240
Db 452 GGGTATATGCCCTTGGATTGTCATGTTCTGACGCCAGACCTTTCACACATTCAGAGACCTT 511
Qy 241 AATAAACAATGCTTAAATGCAACACCGGCTACCCCTCTGTGAACGAGAGTCTTTCGGGTT 300
Db 512 AATAAACAATGCTTAAATGCAACACCGGCTACCCCTCTGTGAACGAGAGTCTTTCGGGTT 571
Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 360
Db 572 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 631
Qy 361 AAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTCGCT 420
Db 632 AAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTCGCT 691
Qy 421 TTTGATTTGAGATCCATGAGAACACACAAAGATTTCTTCACTTACGGGTGTACATG 480
Db 692 TTTGATTTGAGATCCATGAGAACACACAAAGATTTCTTCACTTACGGGTGTACATG 751
Qy 481 TCGGGAAGAAGTTTSSRSAGCTTGGTTTCTTAAAAATTCATCGCGACACATAATGGC 540
Db 752 TCGGGAAGAAGTTTCAAGAGCTTGGTTTCTTAAAAATTCATCGCGACACATAATGGC 811
Qy 541 AAATCGGGGGCCAGAGCAAACTGCAGCAAGCTTTGGAGGTAGTCCAGCAACGATCAAC 600
Db 812 AAATCGGGGGCCAGAGCAAACTGCAGCAAGCTTTGGAGGTAGTCCAGCAACGATCAAC 871
Qy 601 GAGGTCTTCAGGTGCACGGGCCGAGAGCATCTCTCTCTTACAAAATCTGATGGTT 660
Db 872 GAGGTCTTCAGGTGCACGGGCCGAGAGCATCTCTCTCTTACAAAATCTGATGGTT 931

Qy 661 TGTGGCTTCTCTATTTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGTCACACCAA 720
Db 932 TGTGGCTTCTCTATTTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGTCACACCAA 991
Qy 721 AAAATCTGTTTGGTATACAGAGCGCGAGACAGACTCTCCACAAGAGGAATGCCGTCC 780
Db 992 AAAATCTGTTTGGTATACAGAGCGCGAGACAGACTCTCCACAAGAGGAATGCCGTCC 1051
Qy 781 TCGAGGAGAGACTTCCCTGCAAGTCTCAACTTGTAGACCAAAATCTCACCTTGAACCGGG 840
Db 1052 TCGAGGAGAGACTTCCCTGCAAGTCTCAACTTGTAGACCAAAATCTCACCTTGAACCGGG 1111
Qy 841 AAGAAGCTGTGCAGATGCATCCCTCAGCTCGATCCGTTTCCACACTTCCAGGCTTGGCAG 900
Db 1112 AAGAAGCTGTGCAGATGCATCCCTCAGCTCGATCCGTTTCCACACTTCCAGGCTTGGCAG 1171
Qy 901 CTGCTTACCAAGGAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGCAAGAAGGG 960
Db 1172 CTGCTTACCAAGGAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGCAAGAAGGG 1231
Qy 961 AGCACCGACACGACGATTCGAGTTCGAGAAAGGAGCTTTGGAGAAAACAAATTAAGGGCAGT 1020
Db 1232 AGCACCGACACGACGATTCGAGTTCGAGAAAGGAGCTTTGGAGAAAACAAATTAAGGGCAGT 1291
Qy 1021 TGTGACGGCTCTCGCAAGAGAAAGAGAGTGCAAAACATCCCAACGGCGAAGCCCTCC 1080
Db 1292 TGTGACGGCTCTCGCAAGAGAAAGAGAGTGCAAAACATCCCAACGGCGAAGCCCTCC 1351
Qy 1081 GTGACGCGGATCCCAAGTTTACCAGTAGCAGAGAGAGAGCCCACTCACTGCTCCGAGTGC 1140
Db 1352 GTGACGCGGATCCCAAGTTTACCAGTAGCAGAGAGAGAGCCCACTCACTGCTCCGAGTGC 1411
Qy 1141 GGCAGAGCTTTTCAGAACCTACCAACAGCTGTCTTGCACTCCAGGCTCCACAAGAGGAC 1200
Db 1412 GGCAGAGCTTTTCAGAACCTACCAACAGCTGTCTTGCACTCCAGGCTCCACAAGAGGAC 1471
Qy 1201 CGGAGGGCCGGCGGAGTCCGCCACCATGTCTGTGACGGGAGGAGCCGGGAGCGTGT 1260
Db 1472 CGGAGGGCCGGCGGAGTCCGCCACCATGTCTGTGACGGGAGGAGCCGGGAGCGTGT 1531
Qy 1261 TCTCTGACCTCGCGGCGCTCTGGATGAAATCGAGCCGTGGATCGAGGGGAGGAGTGT 1320
Db 1532 TCTCTGACCTCGCGGCGCTCTGGATGAAATCGAGCCGTGGATCGAGGGGAGGAGTGT 1591
Qy 1321 TCTGAGACGGATCTGAGGATGGCTTCCCAAGGAAATCCATCTGGATGAAATGATGAT 1380
Db 1592 TCTGAGACGGATCTGAGGATGGCTTCCCAAGGAAATCCATCTGGATGAAATGATGAT 1651
Qy 1381 GGAGGAAAAATPAAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGGAAAGTTTTC 1440
Db 1652 GGAGGAAAAATPAAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGGAAAGTTTTC 1711
Qy 1441 CGTTCAAATTTATCTCAATATTTCACTCTGAGACGCATACAGGTGAAAAAACCATACAAA 1500
Db 1712 CGTTCAAATTTATCTCAATATTTCACTCTGAGACGCATACAGGTGAAAAAACCATACAAA 1771
Qy 1501 TGTCAATTTTGTGAATATGCTGACGCCAGAGACATCTCTGAGGTATCATCTTGGAGAGA 1560
Db 1772 TGTGAATTTTGTGAATATGCTGACGCCAGAGACATCTCTGAGGTATCATCTTGGAGAGA 1831
Qy 1561 CATCAAGGAAAAACAAACCGATTTGCTGCTGAAGTCAAGAACGATGGTAAATAATCAG 1620
Db 1832 CATCAAGGAAAAACAAACCGATTTGCTGCTGAAGTCAAGAACGATGGTAAATAATCAG 1891
Qy 1621 GACACTGAGATGACATTTAAACCGCTGACAGTGGGCAAAACCAAAATTTGAAAGATTT 1680
Db 1892 GACACTGAGATGACATTTAAACCGCTGACAGTGGGCAAAACCAAAATTTGAAAGATTT 1951
Qy 1681 TTTGATGTGCCAAAGATGTTTACAGGAGTCCACTCTGCAAGAGAGCTTTAAGAGATGCT 1740
Db 1952 TTTGATGTGCCAAAGATGTTTACAGGAGTCCACTCTGCAAGAGAGCTTTAAGAGATGCT 2011
Qy 1741 TCTGTTTTCAGAAATGTTTCTGGGCGAGCGCTGTCTCTCACCAGCACACAAAGATATCAG 1800

Db 2012 TCTGTTTTTCAGAAATGTTCTGGGAGCGCTGTCCTCTCACGAGCACAAAGATACTCAG 2071
Qy 1801 GATTTCCTAATAAATGCACTGATGACAGTGTGTGATAAAGTGAATAAAGAACCCCTACCCCT 1860
Db 2072 GATTTCCTAATAAATGCACTGATGACAGTGTGTGATAAAGTGAATAAAGAACCCCTACCCCT 2131
Qy 1861 GCTTACCTGGACCTGTTTAAAGAGATCAGCAGTTGAATCTCAGGCAAAATCACTCATC 1920
Db 2132 GCTTACCTGGACCTGTTTAAAGAGATCAGCAGTTGAATCTCAGGCAAAATCACTCATC 2191
Qy 1921 TGTAGAACCAAGCGCGATGTTACTCCTCTCCGATGGCAGTACCAACCCATAACCTTGAA 1980
Db 2192 TGTAGAACCAAGCGCGATGTTACTCCTCTCCGATGGCAGTACCAACCCATAACCTTGAA 2251
Qy 1981 GTTAGCCCAAGAGAGAACCAAGAGACCGAGCTGACTGCAGATACAGGCCAAGTGTG 2040
Db 2252 GTTAGCCCAAGAGAGAACCAAGAGACCGAGCTGACTGCAGATACAGGCCAAGTGTG 2311
Qy 2041 GATTGTACGAAACCTTTAAATTTATCCGTGGGGCTCTTCACAAATGCGCGCAATT 2100
Db 2312 GATTGTACGAAACCTTTAAATTTATCCGTGGGGCTCTTCACAAATGCGCGCAATT 2371
Qy 2101 TCTTTGAGTAAAGTTTGATTCCAAGTATCACCTGTCCATTTGTGACCTTCAAGACATTT 2160
Db 2372 TCTTTGAGTAAAGTTTGATTCCAAGTATCACCTGTCCATTTGTGACCTTCAAGACATTT 2431
Qy 2161 TATCCAGAAGTTTAAATGATCACCAGAGACTGGAGCATAAATCAATCTTGACCTTCAT 2220
Db 2432 TATCCAGAAGTTTAAATGATCACCAGAGACTGGAGCATAAATCAATCTTGACCTTCAT 2491
Qy 2221 AAAAATCTCGAACAAGTCTTGTGTAGAGTCGACGTACCGGATGCCCGCAGCGTTG 2280
Db 2492 AAAAATCTCGAACAAGTCTTGTGTAGAGTCGACGTACCGGATGCCCGCAGCGTTG 2551
Qy 2281 CTGGGAAAGATGTCCTCCCTCTCTAGTTTCTGTAACCCCAAGCCCAAGTCTGCTTTC 2340
Db 2552 CTGGGAAAGATGTCCTCCCTCTCTAGTTTCTGTAACCCCAAGCCCAAGTCTGCTTTC 2611
Qy 2341 CCGGCGCAGTCCAAATCCCTGCAATCTGCGAAGGGGAGAGAGCCCTCTTGGGCGAGGC 2400
Db 2612 CCGGCGCAGTCCAAATCCCTGCAATCTGCGAAGGGGAGAGAGCCCTCTTGGGCGAGGC 2671
Qy 2401 AAGGCCCTCTGACTTCAGGATAGACTCTAGCACTTTAGCCCCAAGTAACTGAAGTCC 2460
Db 2672 AAGGCCCTCTGACTTCAGGATAGACTCTAGCACTTTAGCCCCAAGTAACTGAAGTCC 2731
Qy 2461 CACAGCACAGCAGCAATGTGGGGTCCAAAGGGCGCCACCCAGGCAACGCAATCTCAG 2520
Db 2732 CACAGCACAGCAGCAATGTGGGGTCCAAAGGGCGCCACCCAGGCAACGCAATCTCAG 2791
Qy 2521 ATGTTTCTTAAACCAAGTGTTCCTCTGCAACCGGATAGACAAAGAACCCGAGACAAA 2580
Db 2792 ATGTTTCTTAAACCAAGTGTTCCTCTGCAACCGGATAGACAAAGAACCCGAGACAAA 2851
Qy 2581 TTGAAACCTTTCAGTAGTCTTCTCAGGCCACCTCTCGGAGCAGTAAATCAATGGT 2640
Db 2852 TTGAAACCTTTCAGTAGTCTTCTCAGGCCACCTCTCGGAGCAGTAAATCAATGGT 2911
Qy 2641 TCCATCGACTACCCCGCCAGAACGACAGCCGCTGGGCACTCCCGGAGAGACTATTTC 2700
Db 2912 TCCATCGACTACCCCGCCAGAACGACAGCCGCTGGGCACTCCCGGAGAGACTATTTC 2971
Qy 2701 TGTAAATCGGAGTCCAGCAATACTGCAGCAGAAATTTGGTGAAGCCCTTCCAAAAGAGCTG 2760
Db 2972 TGTAAATCGGAGTCCAGCAATACTGCAGCAGAAATTTGGTGAAGCCCTTCCAAAAGAGCTG 3031
Qy 2761 AAGTCCAGCGTGTGGTTCCTTGAAGTGTGACAGCCCGGGGCAATTAAGAGAGGCTAT 2820
Db 3032 AAGTCCAGCGTGTGGTTCCTTGAAGTGTGACAGCCCGGGGCAATTAAGAGAGGCTAT 3091
Qy 2821 GACCTTCCAGTACCATATGTTGAGAGGCATCATCACTGTTTACCGCAGGACTGTGTG 2880

Db 3092 GACCTTCCCAAGTACCATATGTTGTCAGAGGCATCACATCACTGTTACCGCAGGACTGTGTG 3151
Qy 2881 TATCGTTCGAGCGCTGCCTCCCAAAACCAAGGTTCTGAGCTCCAGCGAGGTCCATTC 2940
Db 3152 TATCGTTCGAGCGCTGCCTCCCAAAACCAAGGTTCTGAGCTCCAGCGAGGTCCATTC 3211
Qy 2941 CCAATGTGCTGACTGTTTCAAGCCCTATGTTGCTCCGGGCCACTTTACACTTGTGTG 3000
Db 3212 CCAATGTGCTGACTGTTTCAAGCCCTATGTTGCTCCGGGCCACTTTACACTTGTGTG 3271
Qy 3001 CCTGCTGTAGTCCAGCATCCAGCTCCGACGTTAG ----- 3034
Db 3272 CCTGCTGTAGTCCAGCATCCAGCTCCGACGTTAGAGGAAAGAGCCCTGTGTATCAATCA 3331
Qy 3035 ----- 3034
Db 3332 CACTTATCTAACAGCATGGCACAAAGAGAAACTATGAGAAATTTTATGGGAATGCACAT 3391
Qy 3035 -----AAGGCTCTTGGTGG 3047
Db 3392 TATCGACCAATGACAAAGAACTTGATTCCTAATTAGGGGGGAAAGAGGTCCTTGGTGG 3451
Qy 3048 ATGTCAGTGTCTTACTCCCATGAAATTTTACTTTTACTTTTGAAGCGAATGGT 3107
Db 3452 ATGTCAGTGTCTTACTCCCATGAAATTTTACTTTTACTTTTGAAGCGAATGGT 3511
Qy 3108 GAAAGCTACTGAAATAAGCTGTGATTTGATACATAAAAAATATGAGGAATCTGCAAG 3167
Db 3512 GAAAGCTACTGAAATAAGCTGTGATTTGATTTACTGTATATAAATATGAGGAATCTGCAAG 3571
Qy 3168 GAAACTACAGTTGTGTA 3186
Db 3572 GAACTACAGTTGTGTA 3590

RESULT 8

AAK81093

ID AAK81093 standard; DNA; 14906 BP.

AC AAK81093;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35905.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227109P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 12-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241828P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0256717P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPT; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 35905; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic

CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX Sequence 14906 BP; 3959 A; 3088 C; 3333 G; 4526 T; 0 U; 0 Other;
SQ Query Match 49.0%; Score 1561.2; DB 4; Length 14906;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1482 AGGTGAAACCAATCAATGTAATTTTGTGAATGCTGAGCCCAAGACATCTCT 1541
DB 4689 AGGTGAAACCAATCAATGTAATTTTGTGAATGCTGAGCCCAAGACATCTCT 4748
QY 1542 GAGGTATCACTGGAGAGACATCAAGGAAACCAACCCGATGTTGCTGTAAGTCAA 1601
DB 4749 GAGGTATCACTGGAGAGACATCAAGGAAACCAACCCGATGTTGCTGTAAGTCAA 4808
QY 1602 GAACGATGTTAAATTCAGGACATCTGAAGATGCACTATTAAACCGCTGACAGTGCACAAAC 1661
DB 4809 GAACGATGTTAAATTCAGGACATCTGAAGATGCACTATTAAACCGCTGACAGTGCACAAAC 4868
QY 1662 CAAATTTGAAAGATTTTGTGAGTGGTGCACAAAGATGTTACAGGAGTGCACCTGCAAA 1721
DB 4869 CAAATTTGAAAGATTTTGTGAGTGGTGCACAAAGATGTTACAGGAGTGCACCTGCAAA 4928
QY 1722 GCAGCTTAAGGATGCTCTCTGTTTTCAGAAATGTTCTGGGACGGCTGCTCTCACC 1781
DB 4929 GCAGCTTAAGGATGCTCTCTGTTTTCAGAAATGTTCTGGGACGGCTGCTCTCACC 4988
QY 1782 AGCACACAAAGATATCTCAGGATTTCCATAAAATGCAGCTGATGACAGTGTGATAAAGT 1841
DB 4989 AGCACACAAAGATATCTCAGGATTTCCATAAAATGCAGCTGATGACAGTGTGATAAAGT 5048
QY 1842 GAATTAACCCCTACCCCTGCTTACCTGCACTGTTAAATAAGAGATCAGCAAGTGAAC 1901
DB 5049 GAATTAACCCCTACCCCTGCTTACCTGCACTGTTAAATAAGAGATCAGCAAGTGAAC 5108
QY 1902 TCAGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 1961
DB 5109 TCAGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 5168
QY 1962 TACCACCATTAACCTTTGAAGTTAGCCCAAGAGAGCAACCGAGACCGAGCTGACTG 2021
DB 5169 TACCACCATTAACCTTTGAAGTTAGCCCAAGAGAGCAACCGAGACCGAGCTGACTG 5228
QY 2022 CAGATACAGGCCAAGTGTGATGTCACGAAACCTTTAAATTTATCCGTGGGGCTCT 2081
DB 5229 CAGATACAGGCCAAGTGTGATGTCACGAAACCTTTAAATTTATCCGTGGGGCTCT 5288
QY 2082 TCACAAATTCGCGGCAATTTCTTGAGTAAAGTTGATTCAGATATCACTGTCCATTT 2141
DB 5289 TCACAAATTCGCGGCAATTTCTTTGAGTAAAGTTGATTCAGATATCACTGTCCATTT 5348
QY 2142 TTGTACCTTCAAGACATTTATCCAGAGATTTTAAATGATGACACAGAGACTGGACATAA 2201
DB 5349 TTGTACCTTCAAGACATTTATCCAGAGATTTTAAATGATGACACAGAGACTGGACATAA 5408
QY 2202 ATACAATCTGAGCTTCAATAAACTGTCGAAACCAAGTCTTTGCTTAGAAGTGCAGTAC 2261
DB 5409 ATACAATCTGAGCTTCAATAAACTGTCGAAACCAAGTCTTTGCTTAGAAGTGCAGTAC 5468
QY 2262 CGGATGCGCGCAGGCTGCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 2321
DB 5469 CGGATGCGCGCAGGCTGCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 5528
QY 2322 CAAAGCCCAAGTCTGTTTCCGCGCAGTCCAAATCCCTGCCATCTGCGAGGGGAAGCA 2381
DB 5529 CAAAGCCCAAGTCTGTTTCCGCGCAGTCCAAATCCCTGCCATCTGCGAGGGGAAGCA 5588
QY 2382 GAGCCCTCTCGGCGCAGGCAAGCCCTCTCTGACTTTCAGGGATAGACTCTAGCATTAGC 2441
DB 5589 GAGCCCTCTCGGCGCAGGCAAGCCCTCTCTGACTTTCAGGGATAGACTCTAGCATTAGC 5648

QY 2442 CCCAGTAACCTGAAGTCCCAAGACACAGACAGCAATGTGGGGTCCAGGGGCGCCAC 2501
DB 5649 CCCAGTAACCTGAAGTCCCAAGACACAGACAGCAATGTGGGGTCCAGGGGCGCCAC 5708
QY 2502 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCGACCGGATAAGAC 2561
DB 5709 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCGACCGGATAAGAC 5768
QY 2562 AAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGTCTCTTCTCAGCCCACTCCG 2621
DB 5769 AAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGTCTCTTCTCAGCCCACTCCG 5828
QY 2622 CAGCAGTAACATCAATGTTCCATCGACTACCCCGCCAGACAGACAGCCCTGGGCACC 2681
DB 5829 CAGCAGTAACATCAATGTTCCATCGACTACCCCGCCAGACAGACAGCCCTGGGCACC 5888
QY 2682 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCACGAATACTCGACAGCAATTTGGTGA 2741
DB 5889 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCACGAATACTCGACAGCAATTTGGTGA 5948
QY 2742 GCGCCCTTCCAAAAGACTGAAAGTCCAGCGTGGTGGCTTGGACCTTGACCTTGACCGCCGGGC 2801
DB 5949 GCGCCCTTCCAAAAGACTGAAAGTCCAGCGTGGTGGCTTGGACCTTGACCGCCGGGC 6008
QY 2802 CAATTTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTTCCAGAGGCATCACATCACT 2861
DB 6009 CAATTTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTTCCAGAGGCATCACATCACT 6068
QY 2862 GTTACCGCAGGACTGTGTATTCCTCGCAGCGCTGCTCCCAACCAAGGTTCTCTGAG 2921
DB 6069 GTTACCGCAGGACTGTGTATTCCTCGCAGCGCTGCTCCCAACCAAGGTTCTCTGAG 6128
QY 2922 CTCCAGCAGGCTCGATTTCTCCAAATGTCTGACTTTCCAGAGCCCTATGTTGGTCCCG 2981
DB 6129 CTCCAGCAGGCTCGATTTCTCCAAATGTCTGACTTTCCAGAGCCCTATGTTGGTCCCG 6188
QY 2982 GCCATTTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3041
DB 6189 GCCATTTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6248
QY 3042 TGGTGGATGTTCAGT 3055
DB 6249 TGCATGAGGGGCGT 6262

RESULT 9
ACN44986
ID ACN44986 standard; DNA; 36022 BP.
XX ACN44986;
AC ACN44986;
XX 18-NOV-2004 (first entry)
XX Human genomic sequence hCG37127.
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX Homo sapiens.
XX WO2003073826-A2.
XX 12-SEP-2003.
XX 28-FEB-2003; 2003WO-US006235.
XX 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
XX WPI; 2003-328604/31.
XX

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 1708; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published

SQ Sequence 36022 BP; 9645 A; 7326 C; 8145 G; 10906 T; 0 U; 0 Other;

Query Match 49.0%; Score 1561.2; DB 11; Length 36022;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

| | | | |
|----|-------|--|-------|
| QY | 1482 | AGGTGAAACCACATCAATGTAATTTGTAATATGCTGCAGCCCAAGAGATCTCT | 1541 |
| DB | 15816 | AGGTGAAACCACATCAATGTAATTTGTAATATGCTGCAGCCCAAGAGATCTCT | 15875 |
| QY | 1542 | CAGGTATCACCTTGGAGAGACATCAAGAGAAAAACAAACCGATGTTGCTGCTGAAGTCAA | 1601 |
| DB | 15876 | GAGGTATCACCTTGGAGAGACATCAAGAGAAAAACAAACCGATGTTGCTGCTGAAGTCAA | 15935 |
| QY | 1602 | GAAACGATGTAATAATCAGGACATCAAGAGATGCACTATTAAACCGCTGACGTGCGCAAC | 1661 |
| DB | 15936 | GAAACGATGTAATAATCAGGACATCAAGAGATGCACTATTAAACCGCTGACGTGCGCAAC | 15995 |
| QY | 1662 | CAAAAATTTGAAAGATTTTTTGATGTCGCAAGATGTTACAGCGATCCACCTGGCAA | 1721 |
| DB | 15996 | CAAAAATTTGAAAGATTTTTTGATGTCGCAAGATGTTACAGCGATCCACCTGGCAA | 16055 |
| QY | 1722 | CGAGCTTAAGAGATGCTCTTCTGTTTTTCAGAAATGTTCTGGGCGAGCGTGTCTCTCAC | 1781 |
| DB | 16056 | CGAGCTTAAGAGATGCTCTTCTGTTTTTCAGAAATGTTCTGGGCGAGCGTGTCTCTCAC | 16115 |
| QY | 1782 | AGCACAAAGATCTCAGGATTTCCATAAAATGCACTGATGACAGTGTGATATAAGT | 1841 |
| DB | 16116 | AGCACAAAGATCTCAGGATTTCCATAAAATGCACTGATGACAGTGTGATATAAGT | 16175 |
| QY | 1842 | GAAATAAAACCTACCCCTGCTTACCTGACCTGTTAAAGAGATCAGCAGTTGAAAC | 1901 |
| DB | 16176 | GAAATAAAACCTACCCCTGCTTACCTGACCTGTTAAAGAGATCAGCAGTTGAAAC | 16235 |
| QY | 1902 | TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG | 1961 |
| DB | 16236 | TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG | 16295 |
| QY | 1962 | TACCACCATTAACCTTGAAGTTAGCCCCAAGAGAAAGCAACCGGAGACCGCAGTGA | 2021 |
| DB | 16296 | TACCACCATTAACCTTGAAGTTAGCCCCAAGAGAAAGCAACCGGAGACCGCAGTGA | 16355 |
| QY | 2022 | CAGATACAGGCCAGTGTGATGTCAGCAAAACCTTTAAATTTATCCGTGGGGCTCT | 2081 |
| DB | 16356 | CAGATACAGGCCAGTGTGATGTCAGCAAAACCTTTAAATTTATCCGTGGGGCTCT | 16415 |
| QY | 2082 | TCACAAATTTGCCGGCAATTTCTTTCAGTAAAGTTTGATTCCAAGTATCACCTGTCCAT | 2141 |
| DB | 16416 | TCACAAATTTGCCGGCAATTTCTTTCAGTAAAGTTTGATTCCAAGTATCACCTGTCCAT | 16475 |
| QY | 2142 | TGTACCTTCAAGACATTTTATFCCAGAGTTTAAATGATGACACAGAGCTGGAGCATAA | 2201 |
| DB | 16476 | TGTACCTTCAAGACATTTTATFCCAGAGTTTAAATGATGACACAGAGCTGGAGCATAA | 16535 |

RESULT 10
ACN44985

ID ACN44985 standard; cDNA; 3016 BP.

XX ACN44985;

XX AC

XX 18-NOV-2004 (first entry)

XX Mouse mRNA sequence mC76076.

XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ds.

| | | | |
|----|-------|---|-------|
| QY | 2202 | ATACAACTCTGACGTTTCATAAAAAAGTTCGAAAAAAGTCTCTTGTCTTAGAAGTCGACGTAC | 2261 |
| DB | 16536 | ATACAACTCTGACGTTTCATAAAAAAGTTCGAAAAAAGTCTCTTGTCTTAGAAGTCGACGTAC | 16595 |
| QY | 2262 | CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCCTCCCTCTCTCTAGTTTCTGTAAACC | 2321 |
| DB | 16596 | CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCCCTCCCTCTCTCTAGTTTCTGTAAACC | 16655 |
| QY | 2322 | CAAGCCCAAGTCTGTTTCCCGGGCGCAGTCCAAATCCCTGCGCATCTCGAAGGGGAAGCA | 2381 |
| DB | 16656 | CAAGCCCAAGTCTGTTTCCCGGGCGCAGTCCAAATCCCTGCGCATCTCGAAGGGGAAGCA | 16715 |
| QY | 2382 | GAGCCCTCTCTGGGCGCAGCAAGGCGCCTCTGACTTTCAGGAGATAGACTCTAGCACTTTAGC | 2441 |
| DB | 16716 | GAGCCCTCTCTGGGCGCAGCAAGGCGCCTCTGACTTTCAGGAGATAGACTCTAGCACTTTAGC | 16775 |
| QY | 2442 | CCCAAGTAACCTGTAAGTCCCAACAGACACAGCAAGATGTGGGGTCCAAAGGGCGCCAC | 2501 |
| DB | 16776 | CCCAAGTAACCTGTAAGTCCCAACAGACACAGCAAGATGTGGGGTCCAAAGGGCGCCAC | 16835 |
| QY | 2502 | CAGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGCGCATAGAC | 2561 |
| DB | 16836 | CAGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGCGCATAGAC | 16895 |
| QY | 2562 | AAAAAGACCCGAGACAAAATTTGAAAACCTTCTCCAGTAGTCTCCTTCTCAGCCCAACCTCGG | 2621 |
| DB | 16896 | AAAAAGACCCGAGACAAAATTTGAAAACCTTCTCCAGTAGTCTCCTTCTCAGCCCAACCTCGG | 16955 |
| QY | 2622 | CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCAAGAAACAGACCCCGTGGGCACC | 2681 |
| DB | 16956 | CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCAAGAAACAGACCCCGTGGGCACC | 17015 |
| QY | 2682 | TCCGGGAGAGACTATTTCTGTAATCGAGTGCAGCAATACTCGACAGAAATTTGGTGA | 2741 |
| DB | 17016 | TCCGGGAGAGACTATTTCTGTAATCGAGTGCAGCAATACTCGACAGAAATTTGGTGA | 17075 |
| QY | 2742 | GCCCTCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCCCTTGACGTTGACCAAGCCCGGGC | 2801 |
| DB | 17076 | GCCCTCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCCCTTGACGTTGACCAAGCCCGGGC | 17135 |
| QY | 2802 | CAATTACAGAAAGAGCTATGACCTTCCCAAGTACATATGTCAGAGCATCATCACT | 2861 |
| DB | 17136 | CAATTACAGAAAGAGCTATGACCTTCCCAAGTACATATGTCAGAGCATCATCACT | 17195 |
| QY | 2862 | GTTACCCGAGACTGTGTATCCGTGCGAGCGCTGCCTCCCAAAACCAAGGTTCCCTGAG | 2921 |
| DB | 17196 | GTTACCCGAGACTGTGTGTATCCGTGCGAGCGCTGCCTCCCAAAACCAAGGTTCCCTGAG | 17255 |
| QY | 2922 | CTCCAGCGAGGTCGATTTCTCCAAATGTGCTGACTGTTCCAGAGCCCTATGTTGGCTCCGG | 2981 |
| DB | 17256 | CTCCAGCGAGGTCGATTTCTCCAAATGTGCTGACTGTTCCAGAGCCCTATGTTGGCTCCGG | 17315 |
| QY | 2982 | GCCACTTTTACACTTGTGTGCTGCTGTGTAGTCCAGCATCCAGCTTCGAGCTTAGAAGTCT | 3041 |
| DB | 17316 | GCCACTTTTACACTTGTGTGCTGCTGTGTAGTCCAGCATCCAGCTTCGAGCTTAGAAGTCT | 17375 |
| QY | 3042 | TGGTGGATGTCACT 3055 | |
| DB | 17376 | TGCATGAGGGCGCT 17389 | |


```
Qy 1534 ACATCTCTGAGGTATCACTTGGAGAGACATCAACGAGGAAAAACAAAC---CGATGTTGCT 1590
Dy |||||
Dy 1555 ACATCTCTGAGGTACCACTTGGAGAGACATCAACGAGAGACAGCCGGTGGATGCTGCC 1614
Qy 1591 GCTGAAGTCAAGAACGATGTAAGTAAATAACAGACACTGAAGATGCACTATTA---CCGCT 1647
Dy |||||
Dy 1615 GCTGAGTCCAAAGTGAAGCGCGAGCCAGGAGCGCAGGATGCGCTACTAACGCGTCT 1674
Qy 1648 GACAGTGGCAAAACCAAAATTTGAAGAATTTTGAAGTGTGCAAGATGTTACAGGC 1707
Dy |||||
Dy 1675 GACAGTGGCAAGACCAAAATTTGAAGAATTTTGAAGTGTGCAAGATGTTACAGGC 1734
Qy 1708 AGTCACCTGCAAGACAGCTTAAAGAGATGCTTCTGTTTTTCAAGATGTTTCTGGGAGC 1767
Dy |||||
Dy 1735 AGCCCACTGCAAGACAGCTTAAAGAGATGCTTCTGTTTCCAGAGTGTTC----- 1786
Qy 1768 GCTGTCTCTCACGAGCACAAAGATACTCAGATTTTCCATAAATAATGCAAGCTGATGAC 1827
Dy |||||
Dy 1787 ---TCTCACGAGCACAGCAAGATACTCAGATTTTCCATAAATAATGCAAGCTGAT--- 1839
Qy 1828 AGTCTGATTAAGTGAATAAAGACCTTACCTGCTTACCTGGACCTGTTAAAGAGAGA 1887
Dy |||||
Dy 1840 AGTCTGAGAAGCGAGAGAGCCCTGCCCTTACTTATCTGGACATGCAAGAGAGA-- 1897
Qy 1888 TCAGCAGTTGAAACTCAGGCAAAATAACCTCATCTGTAGAACCAAGCGCGATGTTACTCT 1947
Dy |||||
Dy 1998 -AAGCAGGGAGCTTCAGCCAGCAGCCTCTCTGACACTAGAGGGGTTGGTCTCTTA 1956
Qy 1948 CCTCGGATGCGAGTACCACCCATTAACCTTGAAGTTAGCCCCAAAGAGAAACGAG 2007
Dy |||||
Dy 1957 GCAGGGAGCTGG-----CCATAGGGAGAGATGAT 1989
Qy 2008 ACCGAGCTGACTGACAGATACAGGCCAAGTGTGGATTGTACGAAAAACCTTTAAATTTA 2067
Dy |||||
Dy 1990 CAGGATGCTGACTACAGACATTAAGCCCGTGTGCTGCTCCAGGACAGCCCTTTGAATCTA 2049
Qy 2068 TCCGTGGGGGCTTTCACAAATGCCCCGCAATTTCTTTGAGTAAAGTTTGATTTCCAAAGT 2127
Dy |||||
Dy 2050 TCCCTTGGGCGCTCCAGCGCTGCTGCAATCTCTTGAGCAAGTGTCTGATCCCCAGC 2109
Qy 2128 ATCACTGTCCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTTATGATGACACAG 2187
Dy |||||
Dy 2110 ATTGCTGCCCTTTTGTACTTCAAGACCTTTTATCGGAAGTCTTTATGATGACACAG 2169
Qy 2188 AGACTGGAGCATTAATCAATCCTGACGTTTATAAACTGTGCAAAAGTCTTCTGCTT 2247
Dy |||||
Dy 2170 AGACTTGAGCAGAGTACAACCCCTGACCCGACAAAGAACGGCAGCAAGTCTGTGCTG 2229
Qy 2248 AGAAGTCGACGTACCGGATGCCCGCAGCGTTGCTGGGAAAAAGATGTGCTCCCTCTCT 2307
Dy |||||
Dy 2230 AGGAACAGGGGTACCGGGTGCCCTCCGGTGTGCTGGGAAAAAGATGTCCCTCCCTGTCT 2289
Qy 2308 AGTTTCTGTAAACCAAGCCCAAGTCTGCTTTCCGGCGCAGTCCAAATCCCTGCACTCT 2367
Dy |||||
Dy 2290 GGCCTGCAACAGCCCAAGCCCAAGACTGCTTCTCACACACTCGAAGTCCCTGCACTCA 2349
Qy 2368 GCGAAGGGAGCAGACAGCCCTCTGGGCGCAGCAGCCCTCTGACTTCAGGGATAGAC 2427
Dy |||||
Dy 2350 GAGAAGGCTCGCAGGGGGCTTCGGGGGCAAGCAAGCACCACAGCTTCAGGACACAG 2409
Qy 2428 TCTAGCACTTTAGCCCCCAAGTAACTGAAAGTCCCAAGACACAGACAGATGTGGGGTCT 2487
Dy |||||
Dy 2410 AACAGCACTTTAGCCCCCAAGTAACTGAAAGTCAACAGGTCAACACCAATGCTGGGGGC 2469
Qy 2488 CAAGGGGCGCCACAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCT 2547
Dy |||||
Dy 2470 ACCAG---TGCCACAGGAGCAGCAGTCAAGATGTTTTCCTAAAGTGGCGTCCCTGCT 2526
Qy 2548 GCACCGGATNAGCAAAAGACCCGAGACAAATTTGAACCTTTCCAGTAGCTCTTCT 2607
Dy |||||
Dy 2527 GCTATGGAATGAAGAGACCTGAGCCAAACCTGAAGTCCCTACCAGCCCTCCCGTCT 2586
```

```
Qy 2608 CAGCCCACTCTCGCAGCAGTAACATCAATGGTTTCCATCCGACTACCCGCGCAAGAACAC 2667
Dy |||||
Dy 2587 CAGTCCCCCTTCAGCAGTAATAATAGCAACGGTTCTGTGTAGTATCCCGTGAAGGTTGAC 2646
Qy 2668 AGCCCGTGGGCACTCTCGGGAAGAGACTATTTCTGTAATCGGAGTCCAGCAATACTGCA 2727
Dy |||||
Dy 2647 GGGCCATGGGCACAGCAAGGGAGAGACTACTTACTGCCATCGGAATTTGGGAGTGCAGCA 2706
Qy 2728 GCAGAAATTTGGTGAAGCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTTGACGTT 2787
Dy |||||
Dy 2707 GCAGAGTACAGTGAAGCACATCCCAAAAGACTCAAGTCCAGTGGGTGCTCCCTGGACACA 2766
Qy 2788 GACCAGCCCGGGGCCAATTAACAGAAAGGCTATGACCTTCCCAAGTACCATATGGTCA 2847
Dy |||||
Dy 2767 GAGCATGCAGGAGCAATGCGAAGGGGCTTTGAGCTCCCCAAGTACCATGTGTGTCAGG 2826
Qy 2848 GGCATCATCATCTGTTTACCGCAGACCTGTGTGATCCGTGCGAGGCGCTGCCCTCCCAA 2907
Dy |||||
Dy 2827 AGCATCACTCTTGTGTACCAACAGAGTGTGTGCGCCACCGCTGTCTGCCCAACAA 2886
Qy 2908 CCAAGGTTCTGTAGCTCCAGCGAGTCAATTTCTCCAAATGTGCTGACTGTTTCAGAAAGCCC 2967
Dy |||||
Dy 2887 GCGGTTTCTGAGCCCTGGGAGGTGAGTCAACCCAGTGTGTTGGCTGTGCGAAGGCC 2946
Qy 2968 TATGTTGGTCCCGGCCCATTTTACATTTGTGTGCTGCTGTGTAGTTCAGCATTCAGACTCG 3027
Dy |||||
Dy 2947 TACAGTGCCTCTGAGCCCTGTATACCTGTGAGCCCGTGGGACACACGAGGAGCGCCCA 3006
Qy 3028 ACGTTAGAG 3037
Dy |||||
Dy 3007 GCGCTTGAAG 3016
```

RESULT 11

AAV09023
ID AAV09023 standard; DNA; 10282 BP.

XX AAV09023;

XX AC AAV09023;

XX DT 21-JUL-1998 (first entry)

XX XX Homo sapiens 20q13 amplicon ZABC-1 genomic sequence.

DE 20q13 amplicon; chromosome 20; tumour; detection; ZABC-1 gene;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentata;
KW Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.

XX OS Homo sapiens.

XX PN WO9802539-A1.

XX XX 22-JAN-1998.

XX PF 15-JUL-1997; 97WO-US012343.

XX PR 15-JUL-1996; 96US-00680395.

XX PR 16-OCT-1996; 96US-00731499.

XX PR 17-JAN-1997; 97US-00785532.

XX XX (REGC) UNIV CALIFORNIA.

XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;

XX WPI; 1998-110587/10.

XX DR P-PSDB; AAW23975.

XX New sequences from the 20q13 amplicon - used for detecting chromosomal

XX abnormalities, particularly tumours, and for developing products for

XX treating diseases.

XX Claim 1; Page 64-67; 91pp; English.

XX XX

CC The sequence is that of the genomic sequence of ZABC-1 (zinc finger
CC amplified in breast cancer). It maps to the core of the 20q13.2 amplicon
CC and is overexpressed in primary tumours and breast cancer cell lines
CC having 20q13.2 amplification. The exact coding region for the genomic
CC sequence is not given. The sequence can be used as a probe for the
CC detection of chromosomal abnormalities at 20q13. It and other sequences
CC isolated from the 20q13 amplicon are consistently amplified in primary
CC tumours. These sequences are useful as probes or as probe targets for
CC monitoring the relative copy number of corresponding sequences from a
CC biological sample such as tumour cells. The sequences can also be used in
CC therapeutic applications for modulating the expression of the endogenous
CC gene or the activity of the gene product. Examples of therapeutic
CC approaches include antisense inhibition of gene expression, gene therapy,
CC and monoclonal antibodies that specifically bind the gene products. The
CC products can also be used in the treatment of other diseases, e.g. age-
CC related macular degeneration, Leber's congenital amaurosis and retinitis
CC pigmentation

SQ Sequence 10282 BP; 2820 A; 2222 C; 2191 G; 3047 T; 0 U; 2 Other;

Query Match 46.5%; Score 1482.8; DB 2; Length 10282;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1517; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1482 AGGTGAAAACCATACAAATGTGAATTTGTGAATATCTGTCAGCCAGAGACATCTCT 1541
DB 8204 AGGTGAAAACCATACAAATGTGAATTTGTGAATATCTGTCAGCCAGAGACATCTCT 8263

QY 1542 GAGGTATCACTTGGAGAGACATCAAGGAAAAACAAACCCGATGTGCTGAGTCAA 1601
DB 8264 GAGGTATCACTTGGAGAGACATCAAGGAAAAACAAACCCGATCTTCTGCTCACTCA 8323

QY 1602 GAACGATGGTAAATTCAGGACATCTGAAGATGCACTATTAACCGCTGACAGTGGCAAA 1661
DB 8324 GAACGATGGTAAATTCAGGACATCTGAAGATGCACTATTAACCGCTGACAGTGGCAAA 8383

QY 1662 CAATAATTTGAAAGATTTTGTGATGTCGCAAGATCTTACAGCGATCCACCTGCAAA 1721
DB 8384 CAATAATTTGAAAGATTTTGTGATGTCGCAAGATCTTACAGCGATCCACCTGCAAA 8443

QY 1722 GCAGCTTAAGAGATGCCTCTCTGTTTTTTCAGAAATGTTCTGGCGAGCGTCTCTCACC 1781
DB 8444 CCACCTTAAGAGATGCCTCTCTGTTTTTTCAGAAATGTTCTGGCGAGCGTCTCTCACC 8503

QY 1782 AGCACAAAGATATCTAGGATTTCCATAAAATCAGCTGATGACAGTGTGATAAAGT 1841
DB 8504 ACCACAAAGATATCTAGGATTTCCATAAAATCAGCTGATGACAGTGTGATAAAGT 8563

QY 1842 GAATAAAACCTTACCTCTGATGAAACCAAGCGGATGTTTAAAGAGAGATCAGCAGTTGAAC 1901
DB 8564 GAATAAAACCTTACCTCTGATGAAACCAAGCGGATGTTTAAAGAGAGATCAGCAGTTGAAC 8623

QY 1902 TCAGGCAATTAACCTCATCTGTAGAAACCAAGCGGATGTTTACTCTCTCCGATGGCAG 1961
DB 8624 TCAGGCAATTAACCTCATCTGTAGAAACCAAGCGGATGTTTACTCTCTCCGATGGCAG 8683

QY 1962 TACCACCATTAACCTTGAAGTTAGCCCAAGAGAGAAACCAAGCGGATGTTTACTCTG 2021
DB 8684 TACCACCATTAACCTTGAAGTTAGCCCAAGAGAGAAACCAAGCGGATGTTTACTCTG 8743

QY 2022 CAGATACAGGCCAAGTGTGGATGTGCAGAAAAACCTTTAAATTTATTCGTGGGGCTCT 2081
DB 8744 CAGATACAGGCCAAGTGTGGATGTGCAGAAAAACCTTTAAATAAATTCGTGGGGCTCT 8803

QY 2082 TCACAAATTCGCGGCAATTTCTTTCAGTAAAGTTTGAATTCAGATATCACTGTCCAAT 2141
DB 8804 TCACAAATTCGCGGCAATTTCTTTCAGTAAAGTTTGAATTCAGATATCACTGTCCAAT 8863

QY 2142 TTGTACCTTCAAGACATTTTATTCAGAAAGTTTAAATGATGACACAGAGCTGGAGCATAA 2201
DB 8864 TTGTACCTTCAAGACATTTTATTCAGAAAGTTTAAATGATGACACACACCTCCACATAA 8923

QY 2202 ATACAACTCTGACGTTTCAATAAAACCTGTCGAAACAAAGTCTTGTGTTAGATGTCAGCTAC 2261

DB 8924 ATACAACTCTGACGTTTCAATAAAACTGTGGAACAAAGTCTTGTCTTACACTCGACGTAC 8983
QY 2262 CGGATGCCCGCCAGCGTTGTCTGGGAAAAGATGTGCTCTCCCTCTCTCTAGTTCTTGTAAACC 2321
DB 8984 CCGATGCCCGCCAGCGTTGTCTGGGAAAAGATGTGCTCTCCCTCTCTCTAGTTCTTGTAAACC 9043

QY 2322 CAAGCCCAAGTCTGCTTTCCCGGGCAGTCAAAATCCCTGCGCATCTGCAAGGGGAAGA 2381
DB 9044 CAAGCCCAAGTCTGCTTTCCCGGGCAGTCAAAATCCCTGCGCATCTGCAAGGGGAAGA 9103

QY 2382 GAGCCCTCTCTGGCCAGCAGCAAGGCCCTCTGACATTCAGGGATAGACTCTAGCATTAGC 2441
DB 9104 GAGCCCTCTCTGGCCAGCAGCAAGGCCCTCTGACATTCAGGGATAGACTCTAGCATTAGC 9163

QY 2442 CCCAAGTAACCTGAAGTCCCAAGACCAAGCAAGTGTGGGGTCCAAAGGGGGCCGCCAC 2501
DB 9164 CCCAAGTAACCTGAAGTCCCAAGACCAAGCAAGTGTGGGGTCCAAAGGGGGCCGCCAC 9223

QY 2502 CAGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTGCGTGCACCGGATAAGAC 2561
DB 9224 CAGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTGCGTGCACCGGATAAGAC 9283

QY 2562 AAAAGAGCCCGAGACAAATTTGAAACCTCTTCCAGTAGTCTCTTCTCAGCCACCCCTCGG 2621
DB 9284 AAAAGAGCCCGAGACAAATTTGAAACCTCTTCCAGTAGTCTCTTCTCAGCCACCCCTCGG 9343

QY 2622 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGCACC 2681
DB 9344 CAGCAGTAACATCAATGGTTCCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGCACC 9403

QY 2682 TCCGGGAAGAGACTATTCTGTAATCGGAGTGCAGCAATATCTGAGCAGAGAAATTTGGTGA 2741
DB 9404 TCCGGGAAGAGACTATTCTGTAATCGGAGTGCAGCAATATCTGAGCAGAGAAATTTGGTGA 9463

QY 2742 GCCCTTCCAAAAGAGACTGAAAGTCCAGCGTGTGCTTCCAGTGTGACGTTGACCGCGGGGC 2801
DB 9464 GCCCTTCCAAAAGAGACTGAAAGTCCAGCGTGTGCTTCCAGTGTGACGTTGACCGCGGGGC 9523

QY 2802 CAATTACAGAAAGGCTATGACCTTCCCAAGTACCATATGTTGTCAGAGGCATCACAATCACT 2861
DB 9524 CAATTACAGAAAGGCTATGACCTTCCCAAGTACCATATGTTGTCAGAGGCATCACAATCACT 9583

QY 2862 GTTACCGAGAGCTGTGTGATTCGTCGAGCGCGTGCCTCCCAAAACCAAGGTTCCCTGAG 2921
DB 9584 GTTACCGAGAGCTGTGTGATTCGTCGAGCGCGTGCCTCCCAAAACCAAGGTTCCCTGAG 9643

QY 2922 CTCCAGGAGGTCGATTCTCCAAATGTGCTGACTGTTTCAGAAAGCCCTATGTTGGTCCGG 2981
DB 9644 CTCCAGGAGGTCGATTCTCCAAATGTGCTGACTGTTTCAGAAAGCCCTATGTTGGTCCGG 9703

QY 2982 GCCACTTTTACACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3041
DB 9704 GCCACTTTTACACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9763

QY 3042 TGGTGGATGTCAGT 3055
DB 9764 TGCATGAGGGCGT 9777

RESULT 12

AAV04698
ID AAV04698 standard; cDNA; 1507 BP.

AC AAV04698;

XX 21-JUL-1998 (first entry)

XX Homo sapiens 20q13 amplicon cc49 transcript.

DE 20q13 amplicon; chromosome 20; tumour; detection; C2H2 zinc finger;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentation;

AAK81095
ID AAK81095 standard; DNA; 566 BP.
XX
AC AAK81095;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35907.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 22-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 35907; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients' own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 566 BP; 170 A; 127 C; 134 G; 135 T; 0 U; 0 Other;
XX
XX Query Match 17.7%; Score 563.6; DB 4; Length 566;
XX Best Local Similarity 98.9%; Pred. No. 4.7e-158;
XX Matches 560; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
XX
QY 113 TGTCAATGAAGGCGCGTGTGTTCATTCCGAGCTACACAGAAAATGTCATCC 172
DB 1 TGTCAATGAAGGCGCGTGTGTTCATTCCGAGCTACACAGAAAATGTCATCC 60
XX
QY 173 AAATCGAGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGACCTTCACACATTCAG 232
DB 61 AAATCGAGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGACCTTCACACATTCAG 120
XX
QY 233 AAGACCTTAATAACATGCTTAATGCAACACCGGCTACCTCTGTGAACAGCAGTTC 292
DB 121 AAGACCTTAATAACATGCTTAATGCAACACCGGCTACCTCTGTGAACAGCAGTTC 180
XX
QY 293 TTCGGGTGAAGCAGAGTATCTCAGTCGCTTGATAAAGTCAAGTCGGAACAGACCTC 352
DB 181 TTCGGGTGAAGCAGAGTATCTCAGTCGCTTGATAAAGTCAAGTCGGAACAGACCTC 240
XX
QY 353 CCAAGGAAAAGAAATGCAAGGAAAATGAATTTAGTGTGAGTATGTGGGAGACATTTA 412
DB 241 CCAAGGAAAAGAAATGCAAGGAAAATGAATTTAGTGTGAGTATGTGGGAGACATTTA 300
XX
QY 413 GAGTCGCTTTTGAATGTTGAGATCCATGAGAACACACAAAGATTTCTTCACTTACGGGT 472
DB 301 GAGTCGCTTTTGAATGTTGAGATCCATGAGAACACACAAAGATTTCTTCACTTACGGGT 360
XX
QY 473 GTAACATGTCGGAAGAGAGTTSRSSAGCCCTTGTCTTTAAATACATGTCGGGACAC 532
DB 361 GTAACATGTCGGAAGAGAGTTCAGAGGAGCCCTTGTGTTTCTTAAATACATGTCGGGACAC 420

QY 533 ATAATGGCAAAATCGGGGCGCAGAACAAATCGCAGCAAGGCTTGAGAGTAGTCTCCGCAA 592
DB 421 ATAATGGCAAAATCGGGGCGCAGAACAAATCGCAGCAAGGCTTGAGAGTAGTCTCCGCAA 480
XX
QY 593 CGATCAACGAGGTGCTCCAGGTGCACGGCGGCGAGAGCATCTCTCTTACAAAATCT 652
DB 481 CGATCAACGAGGTGCTCCAGGTGCACGGCGGCGAGAGCATCTCTCTTACAAAATCT 540
XX
QY 653 GCATGTTTGTGGCTTCTCTATTTCOA 678
DB 541 GCATGTTTGTGGCTTCTCTATTTCOA 566
XX

RESULT 15

AAV88042

ID AAV88042 standard; cDNA; 469 BP.

XX

AC AAV88042;

XX

DT 12-FEB-1999 (first entry)

XX

DE EST clone FP233.

XX

KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX

OS Homo sapiens;

XX

PN WO9845437-A2.

XX

PD 15-OCT-1998.

XX

PF 10-APR-1998; 98WO-US0006956.

XX

PR 10-APR-1997; 97US-00837312.

XX

(GENY) GENETICS INST INC.

XX

PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;

PI Spaulding V, Agostino MJ;

XX

WPI; 1999-070078/06.

XX

PT New polynucleotides encoding human secreted proteins - derived from e.g.
PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,
PT pituitary, retina and colon cDNA libraries.

XX

PS Claim 1; Page 263; 641pp; English.

XX

CC The present sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene therapy

XX

SQ Sequence 469 BP; 125 A; 119 C; 132 G; 93 T; 0 U; 0 Other;

XX

Query Match 13.9%; Score 444.4; DB 2; Length 469;

Best Local Similarity 99.8%; Pred. No. 2.9e-122;

Matches 445; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX

QY 1032 CTCGCAAGAGAAAGAGAGTGCAAACTCCACGGCGAGCCCTCCGTGGACGCGA 1091

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:43:42 ; Search time 502.445 Seconds
(without alignments)
10375.608 Million cell updates/sec

Title: US-08-731-499-10
Perfect score: 3186
Sequence: 1 ATGCAATCGAAGTACAGG.....GGAACACTACAGTTGTGTAA 3186

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|----------------------|
| 1 | 3183.6 | 99.9 | 3186 | 4 | US-08-892-695-10 |
| 2 | 3040.6 | 95.4 | 5632 | 3 | US-09-560-594-3 |
| 3 | 3040.6 | 95.4 | 5632 | 4 | US-09-949-016-862 |
| 4 | 3040.6 | 95.4 | 5632 | 4 | US-09-949-016-4262 |
| 5 | 1561.2 | 49.0 | 20022 | 4 | US-09-949-016-12604 |
| 6 | 1561.2 | 49.0 | 20023 | 4 | US-09-949-016-16004 |
| 7 | 1559.6 | 49.0 | 10365 | 4 | US-08-892-695-9 |
| 8 | 1106.2 | 34.7 | 1507 | 2 | US-08-680-395-3 |
| 9 | 1106.2 | 34.7 | 1507 | 4 | US-08-892-695-3 |
| 10 | 600.6 | 18.9 | 601 | 4 | US-09-949-016-34700 |
| 11 | 600.6 | 18.9 | 601 | 4 | US-09-949-016-151880 |
| 12 | 598.2 | 18.8 | 601 | 4 | US-09-949-016-34699 |
| 13 | 598.2 | 18.8 | 601 | 4 | US-09-949-016-151879 |
| 14 | 66.8 | 2.1 | 2765 | 4 | US-09-620-312D-61 |
| 15 | 66.8 | 2.1 | 2799 | 4 | US-09-949-016-4961 |
| 16 | 66.8 | 2.1 | 3039 | 4 | US-09-949-016-1038 |
| 17 | 66.8 | 2.1 | 12323 | 4 | US-09-949-016-16703 |
| 18 | 59.8 | 1.9 | 2785 | 4 | US-09-774-528-105 |
| 19 | 59.2 | 1.9 | 696 | 4 | US-09-451-651-4 |
| 20 | 59.2 | 1.9 | 2920 | 4 | US-09-620-312D-1084 |
| 21 | 57.6 | 1.8 | 2752 | 4 | US-09-949-016-3563 |
| 22 | 57.6 | 1.8 | 18824 | 4 | US-09-949-016-15305 |
| 23 | 56.8 | 1.8 | 936 | 4 | US-09-016-434-312 |
| 24 | 55.2 | 1.7 | 2771 | 4 | US-09-976-594-691 |
| 25 | 55 | 1.7 | 4272 | 4 | US-09-620-312D-586 |
| 26 | 54.8 | 1.7 | 265 | 4 | US-09-016-434-836 |
| 27 | 54.4 | 1.7 | 2489 | 4 | US-09-774-528-280 |

ALIGNMENTS

RESULT 1

US-08-892-695-10
; Sequence 10, Application US/08892695A
; Patent No. 6808878

GENERAL INFORMATION:
; APPLICANT: Gray, Joe W

; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In

; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David

; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES

; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892.695A

; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532

; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499

; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395

; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10

; LENGTH: 3186
; TYPE: DNA

; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: ZABC1 Open
; OTHER INFORMATION: Reading Frame

US-08-892-695-10

Query Match 99.9%; Score 3183.6; DB 4; Length 3186;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| Qy | 1 | ATGCAATCGAAGTACAGGAAACATGCCAACTCAATCCCTCTTAATCTTACATGATGGG | 60 |
| Db | 1 | ATGCAATCGAAGTACAGGAAACATGCCAACTCAATCCCTCTTAAATGATGATGGG | 60 |
| Qy | 61 | CCAGAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG | 120 |
| Db | 61 | CCAGAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG | 120 |
| Qy | 121 | AAAGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCCATCCAAATCGAG | 180 |
| Db | 121 | AAAGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCCATCCAAATCGAG | 180 |
| Qy | 181 | GGGTATATGCCCTTGGATTCATGTTCTGCGAGCCGACCTTCACACATTCAGAGACCTT | 240 |

Db 181 GGGTATATGCCCTTGGATTGCGATGTTCTGCAGCCAGACCTTTACACATTTCAGAAAGACCTT 240
Qy 241 AATAAAATGCTCTTAATCAACACCGGCTTACCCTCTGTGAACACAGCAGTTCTTTCGGGTT 300
Db 241 AATAAAATGCTCTTAATCAACACCGGCTTACCCTCTGTGAACACAGCAGTTCTTTCGGGTT 300
Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGACCTTCCCAAGGAA 360
Db 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGACCTTCCCAAGGAA 360
Qy 361 AAGAAATTCGAAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCGACAGATTTAGAGTGCCT 420
Db 361 AAGAAATTCGAAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCGACAGATTTAGAGTGCCT 420
Qy 421 TTTGATGTTGAGATCCACATCAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 480
Db 421 TTTGATGTTGAGATCCACATCAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 480
Qy 481 TCGGAAGAAAGTTSRRSSAGCCTTGGTTTCTTAAAAATCACATCGCGACACATAATGGC 540
Db 481 TCGGAAGAAAGTTCAGAGGCTTGGTTTCTTAAAAATCACATCGCGACACATAATGGC 540
Qy 541 AAATCGGGGGCCAGAAGCAAACTGCAGCAAGCTTTGGAGAGTAGTTCAGCAAAAGATCAAC 600
Db 541 AAATCGGGGGCCAGAAGCAAACTGCAGCAAGCTTTGGAGAGTAGTTCAGCAAAAGATCAAC 600
Qy 601 GAGTCTGTCAGGTGCGACGCGCGAGAGCATCTCCTCTTACAAATCTGCATGGTT 660
Db 601 GAGTCTGTCAGGTGCGACGCGCGAGAGCATCTCCTCTTACAAATCTGCATGGTT 660
Qy 661 TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAA 720
Db 661 TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAA 720
Qy 721 AAAATGCTGTTTCGATCCAGCAGCGCGCAGACAGACTCTCCATCAAGAGGAAATGCGGTCC 780
Db 721 AAAATGCTGTTTCGATCCAGCAGCGCGCAGACAGACTCTCCATCAAGAGGAAATGCGGTCC 780
Qy 781 TCGAGGAGGAGTCTCTCAGTTGTTCAACTTGGAGACCAAAATCTCACCCCTGAAACGGGG 840
Db 781 TCGAGGAGGAGTCTCTCAGTTGTTCAACTTGGAGACCAAAATCTCACCCCTGAAACGGGG 840
Qy 841 AAGAAGCCTGTGCAGATGCCCTCAGCTCGATCCGTTTCAACACCTTCCAGGCTTTGGCAG 900
Db 841 AAGAAGCCTGTGCAGATGCCCTCAGCTCGATCCGTTTCAACACCTTCCAGGCTTTGGCAG 900
Qy 901 CTGGCTACCAAAGGAAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGGCAAGAAGGG 960
Db 901 CTGGCTACCAAAGGAAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGGCAAGAAGGG 960
Qy 961 AGCACCGAACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATAAGGGCAGT 1020
Db 961 AGCACCGAACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATAAGGGCAGT 1020
Qy 1021 TGTGAGGCGCTCTCCGAAGAGAAAAGAGTGCAAAACACTCCCAAGGAGAGCGCCCTCC 1080
Db 1021 TGTGAGGCGCTCTCCGAAGAGAAAAGAGTGCAAAACACTCCCAAGGAGAGCGCCCTCC 1080
Qy 1081 GTGGACGCGGATCCCAAGTACCCAGTAGCAAGGAGAGAGCCCACTCACTGCTCCGAGTGC 1140
Db 1081 GTGGACGCGGATCCCAAGTACCCAGTAGCAAGGAGAGAGCCCACTCACTGCTCCGAGTGC 1140
Qy 1141 GGCAAGCTTTTCAGAACTTACACAGCTGCTTGCATCTCAGGGTCCACAAGAGGAC 1200
Db 1141 GGCAAGCTTTTCAGAACTTACACAGCTGCTTGCATCTCAGGGTCCACAAGAGGAC 1200
Qy 1201 CGGAGGCGCGGCGGAGTTCGCCACCATGTCTGTGAGCGGAGGAGCGCGGGACGTTGT 1260
Db 1201 CGGAGGCGCGGCGGAGTTCGCCACCATGTCTGTGAGCGGAGGAGCGCGGGACGTTGT 1260
Qy 1261 TCTCTGATCTCGCGCGCCCTCTGATGAATAATGAGCGCGTGGATTCGAGGGGAGGTGTT 1320
Db 1261 TCTCTGATCTCGCGCGCCCTCTGATGAATAATGAGCGCGTGGATTCGAGGGGAGGTGTT 1320

Qy 1321 TCTGAAGACGATCTGAGGATGGCTTCCGAAGGAATCCATCTGGATAAAAAATGATGAT 1380
Db 1321 TCTGAAGACGATCTGAGGATGGCTTCCGAAGGAATCCATCTGGATAAAAAATGATGAT 1380
Qy 1381 GGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAAGTTTTTC 1440
Db 1381 GGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAAGTTTTTC 1440
Qy 1441 CGTTCAAAATTTACCTCAATATTTCACTCAGAACGCATACAGGTGAAGAACCCATACAAA 1500
Db 1441 CGTTCAAAATTTACCTCAATATTTCACTCAGAACGCATACAGGTGAAGAACCCATACAAA 1500
Qy 1501 TGTGAATTTTGTGAATATGCTGCAGCCCGAGAACATCTCTGAGGTATCACTTGGAGAGA 1560
Db 1501 TGTGAATTTTGTGAATATGCTGCAGCCCGAGAACATCTCTGAGGTATCACTTGGAGAGA 1560
Qy 1561 CATCAAGGAAAAAACAACCGATGTTGCTGAAGTCAAGAACGATGGTAAAAATCAG 1620
Db 1561 CATCAAGGAAAAAACAACCGATGTTGCTGAAGTCAAGAACGATGGTAAAAATCAG 1620
Qy 1621 GACACTGAAGTGCACCTATTAAACCGCTGACAGTGCAGAACCCAAAAATTTGAAAAGATTT 1680
Db 1621 GACACTGAAGTGCACCTATTAAACCGCTGACAGTGCAGAACCCAAAAATTTGAAAAGATTT 1680
Qy 1681 TTTGATGCTCCAAAGATGTTACAGGCAGTCCACTGCAAGACAGCTTTAAGGAGATGCCT 1740
Db 1681 TTTGATGCTCCAAAGATGTTACAGGCAGTCCACTGCAAGACAGCTTTAAGGAGATGCCT 1740
Qy 1741 TCTGTTTTTCAAGATGTTCTGGGCGAGCTGCTCTCAACAGCACAAAGATATCTCAG 1800
Db 1741 TCTGTTTTTCAAGATGTTCTGGGCGAGCTGCTCTCAACAGCACAAAGATATCTCAG 1800
Qy 1801 GATTTCCATAAAATGACAGTGTGACAGTCTGATAAAGTGAATAAAAAACCCCTACCCCT 1860
Db 1801 GATTTCCATAAAATGACAGTGTGACAGTCTGATAAAGTGAATAAAAAACCCCTACCCCT 1860
Qy 1861 GCTTACCTGGAACCTGTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAAACCCTCATC 1920
Db 1861 GCTTACCTGGAACCTGTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAAACCCTCATC 1920
Qy 1921 TGTAGAACCAAGGCGGATGTTATCTCTCCGAGTGGCAGTACACCCATTAACCTTGAA 1980
Db 1921 TGTAGAACCAAGGCGGATGTTATCTCTCCGAGTGGCAGTACACCCATTAACCTTGAA 1980
Qy 1981 GTTAGCCCCCAAGAGAGCAAAACGAGACCGCAGCTGACATGACAGTACAGGCCAAGTGTG 2040
Db 1981 GTTAGCCCCCAAGAGAGCAAAACGAGACCGCAGCTGACATGACAGTACAGGCCAAGTGTG 2040
Qy 2041 GATTTGTCACAAAAAACCTTTAAATTTATCCGTGGGGGCTCTTTCACAAATTCGCCGCAATT 2100
Db 2041 GATTTGTCACAAAAAACCTTTAAATTTATCCGTGGGGGCTCTTTCACAAATTCGCCGCAATT 2100
Qy 2101 TCTTTGAGTAAAAGTTGATTCGAAAGTATACCTGTGTCATTTTGTATCCTTCAAGACATTT 2160
Db 2101 TCTTTGAGTAAAAGTTGATTCGAAAGTATACCTGTGTCATTTTGTATCCTTCAAGACATTT 2160
Qy 2161 TATCCAGAGTTTTTAATGATGCACAGAGACTGGAGCATATAATCAATCTCTGAGCTTCAT 2220
Db 2161 TATCCAGAGTTTTTAATGATGCACAGAGACTGGAGCATATAATCAATCTCTGAGCTTCAT 2220
Qy 2221 AAAAATCTGTGCAAAACAAAGTCTTGTAGAAAGTGCAGTACCCGATGCCGCCAGCGTTG 2280
Db 2221 AAAAATCTGTGCAAAACAAAGTCTTGTAGAAAGTGCAGTACCCGATGCCGCCAGCGTTG 2280
Qy 2281 CTGGGAAAAGATGTCCTCTCCCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2340
Db 2281 CTGGGAAAAGATGTCCTCTCCCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2340
Qy 2341 CCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGAGAGAGCCCTCTCTGGGCGAGGC 2400
Db 2341 CCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGAGAGAGCCCTCTCTGGGCGAGGC 2400

| | | | |
|----|------|---|------|
| Qy | 2401 | AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAGTAACTCTGAGTCC | 2460 |
| Db | 2401 | AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAGTAACTCTGAGTCC | 2460 |
| Qy | 2461 | CACAGACCACAGCAGAATGTGGGGTCCAAAGGGCCGCCACACAGGCAACAGCAATCTGAG | 2520 |
| Db | 2461 | CACAGACCACAGCAGAATGTGGGGTCCAAAGGGCCGCCACAGGCAACAGCAATCTGAG | 2520 |
| Qy | 2521 | ATGTTTCTTAAACCCAGTGTTTCCCTTGCAACCGGATAAGAATAAAAGACCCGAGACAAAA | 2580 |
| Db | 2521 | ATGTTTCTTAAACCCAGTGTTTCCCTTGCAACCGGATAAGAATAAAAGACCCGAGACAAAA | 2580 |
| Qy | 2581 | TTGAAACCTCTTCCAGTAGTCTCTTCTCAGCCACCTCGGCAAGCATACATCAATGGT | 2640 |
| Db | 2581 | TTGAAACCTCTTCCAGTAGTCTCTTCTCAGCCACCTCGGCAAGCATACATCAATGGT | 2640 |
| Qy | 2641 | TCCATCGACTACCCCGCAAGAACGACAGCCCGTGGGCACCTCGGGAAGAGACTATTTTC | 2700 |
| Db | 2641 | TCCATCGACTACCCCGCAAGAACGACAGCCCGTGGGCACCTCGGGAAGAGACTATTTTC | 2700 |
| Qy | 2701 | TGTAACTCGAGTGCACGAATATCTGCAGCAGAAATTTGGTGAGGCCCTTCCAAAAAGACTG | 2760 |
| Db | 2701 | TGTAACTCGAGTGCACGAATATCTGCAGCAGAAATTTGGTGAGGCCCTTCCAAAAAGACTG | 2760 |
| Qy | 2761 | AAGTCCAGCGTGGTTCGCCCTTGAGTTCAGCCCGGGGGCCAAATTACAGAAGGCGTAT | 2820 |
| Db | 2761 | AAGTCCAGCGTGGTTCGCCCTTGAGTTCAGCCCGGGGGCCAAATTACAGAAGGCGTAT | 2820 |
| Qy | 2821 | GACCTTCCCAAGTACCATATGTTGAGAGGCATCACATCACTGTTTACCGCAGGACTGTGTG | 2880 |
| Db | 2821 | GACCTTCCCAAGTACCATATGTTGAGAGGCATCACATCACTGTTTACCGCAGGACTGTGTG | 2880 |
| Qy | 2881 | TATCCGTGCGAGGGCGTCCCTCCAAACCAAGGTTCTTGAGCTCCAGCGAGGTCGATTC | 2940 |
| Db | 2881 | TATCCGTGCGAGGGCGTCCCTCCAAACCAAGGTTCTTGAGCTCCAGCGAGGTCGATTC | 2940 |
| Qy | 2941 | CCAAATGTGTGACTGTTTCAGAGGCCCTATGTTGGCTCCGGGCCACATTTACACTTGTGTG | 3000 |
| Db | 2941 | CCAAATGTGTGACTGTTTCAGAGGCCCTATGTTGGCTCCGGGCCACATTTACACTTGTGTG | 3000 |
| Qy | 3001 | CCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTTCTGGTGGATGTGAGTGCTTGA | 3060 |
| Db | 3001 | CCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTTCTGGTGGATGTGAGTGCTTGA | 3060 |
| Qy | 3061 | CTCCCCATGAAATTAAATTTTACTTTCATCCTTTGAGAGGCAATGGTGAAGCTACTGAA | 3120 |
| Db | 3061 | CTCCCCATGAAATTAAATTTTACTTTCATCCTTTGAGAGGCAATGGTGAAGCTACTGAA | 3120 |
| Qy | 3121 | ATAAGCTGTGATTGTACTGTACATAAAAAATATGAGGAATCTCGCAAGGAACACTACAGTT | 3180 |
| Db | 3121 | ATAAGCTGTGATTGTACTGTACATAAAAAATATGAGGAATCTCGCAAGGAACACTACAGTT | 3180 |
| Qy | 3181 | GTGTAA 3186 | |
| Db | 3181 | GTGTAA 3186 | |

```

RESULT 2
US-09-560-594-3
; Sequence 3, Application US/09560594
; Patent No. 6242590
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF ZINC FINGER PROTEIN-217 EXPRESSION
; FILE REFERENCE: RTS-0144
; CURRENT APPLICATION NUMBER: US/09/560,594
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```


QY 3048 ATGTCAGTCTTACTCCCATGAAATTAATTTTACTCTTATCTTCTTGTAGAGCGAATGGT 3107
|||||
Db 3452 ATGTCAGTCTTACTCCCATGAAATTAATTTTACTCTTATCTTGTAGAGCGAATGGT 3511
|||||
QY 3108 GAAAGCTACTGAAATGAAGCTGTGATTGTACTGTACATAAAACATATGAGGAATCTGCAAG 3167
|||||
Db 3512 GAAAGCTACTGAAATGAAGCTGTGATTGTACTGTACATAAAACATATGAGGAATCTGCAAG 3571
|||||
QY 3168 GAACACTACAGTTGTGTAA 3186
|||||
Db 3572 GAACACTACAGTTGTGTAA 3590
|||||

RESULT 3

US-09-949-016-862
; Sequence 862, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 862
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-862

Query Match 95.4%; Score 3040.6; DB 4; Length 5632;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;

QY 1 ATGCAATCGAAGTACAGGAACATGCCAATCAATCCCTCTTAATGTACATGATGGG 60
Db 272 ATGCAATCGAAGTACAGGAACATGCCAATCAATCCCTCTTAATGTACATGATGGG 331
|||||
QY 61 CCAGAAATGATTGGCAGCTCTCTGGCAGTCGATGGAGATGGAGATGCCCTTGTCAATG 120
|||||
Db 332 CCAGAAATGATTGGCAGCTCTCTGGCAGTCGATGGAGATGGAGATGCCCTTGTCAATG 391
|||||
QY 121 AAAGGACCGCTGTTGTTCCATTCGAGCTACACAAAGAAAAATGTCCATCCAAATCGAG 180
Db 392 AAAGGACCGCTGTTGTTCCATTCGAGCTACACAAAGAAAAATGTCCATCCAAATCGAG 451
|||||
QY 181 GGGTATATGCCCTTGGATGATGTTCTGAGCGAGACCTTCACATTCAGAGACCTT 240
Db 452 GGGTATATGCCCTTGGATGATGTTCTGAGCGAGACCTTCACATTCAGAGACCTT 511
|||||
QY 241 AATAAACATGCTTAATCAACACCGGCTACCTCTGTGAACAGAGCTTCTCGGGTT 300
Db 512 AATAAACATGCTTAATCAACACCGGCTACCTCTGTGAACAGAGCTTCTCGGGTT 571
|||||
QY 301 GAAGCAGAGTATCTCAGTCCGCTTGATTAAGTCAAGTGGCAACAGAACCTCCCAAGGAA 360
Db 572 GAAGCAGAGTATCTCAGTCCGCTTGATTAAGTCAAGTGGCAACAGAACCTCCCAAGGAA 631
|||||
QY 361 AAGAAATCGAAGAAATGAATTTAGCTGTGAGGATGTGGGCGACATTTAGAGTCGCT 420
Db 632 AAGAAATCGAAGAAATGAATTTAGCTGTGAGGATGTGGGCGACATTTAGAGTCGCT 691
|||||
QY 421 TTTGATGTTGAGATCCATGAGAACACACAAAGATTTCTTCACTTACGGGTGTAACATG 480
|||||

Db 692 TTTGATGTTGAGATCCATGAGAACACACAAAGATTTCTTCACTTACGGGTGAACATG 751
QY 481 TCGCGAAGAAAGTTTSRRSSAGCTTGGTTTCTTAAAAATCAATCGCGACACATTAATGCG 540
|||||
Db 752 TCGCGAAGAAAGTTTCAAGGAGCTTGGTTTCTTAAAAATCAATCGCGACACATTAATGCG 811
|||||
QY 541 AAATCGGGGGCCAGAAAGCAAACTGCAGCAAGCTTTGGAGAGTAGTCCAGCAACGATCAAC 600
Db 812 AAATCGGGGGCCAGAAAGCAAACTGCAGCAAGCTTTGGAGAGTAGTCCAGCAACGATCAAC 871
|||||
QY 601 GAGGTGCTCCAGGTGCACGCGCCGAGAGCATCTCTCTCTTCAAAAATCTGCAATGTT 660
Db 872 GAGGTGCTCCAGGTGCACGCGCCGAGAGCATCTCTCTCTTCAAAAATCTGCAATGTT 931
|||||
QY 661 TGTGCTTCTTATTTTCAAATAAAGAAAGTCTAATGAGACCGCAAGGTGACACCAAAA 720
Db 932 TGTGCTTCTTATTTTCAAATAAAGAAAGTCTAATGAGACCGCAAGGTGACACCAAAA 991
|||||
QY 721 AAAACTGCTTTCCGTTACCAGAGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCC 780
Db 992 AAAACTGCTTTCCGTTACCAGAGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTCC 1051
|||||
QY 781 TCGAGGGAGGACTTCTCTGAGTTGTTCAACTTGGAGACCAAAAATCTCACCTGAAACGGG 840
Db 1052 TCGAGGGAGGACTTCTCTGAGTTGTTCAACTTGGAGACCAAAAATCTCACCTGAAACGGG 1111
|||||
QY 841 AAGAAAGCTGTGAGATGATCTCTGAGTCTGATCTGATCTTCAACCTTCCAGGCTTGGCAG 900
Db 1112 AAGAAAGCTGTGAGATGATCTCTGAGTCTGATCTGATCTTCAACCTTCCAGGCTTGGCAG 1171
|||||
QY 901 CTGGCTACCAAGGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGGCAAGAGGG 960
Db 1172 CTGGCTACCAAGGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGGCAAGAGGG 1231
|||||
QY 961 AGCAACGACAAACGAGTTCGAGTTCCGAGAGGAGCTTGGAGAAAACAAATAAGGCGAGT 1020
Db 1232 AGCAACGACAAACGAGTTCGAGTTCCGAGAGGAGCTTGGAGAAAACAAATAAGGCGAGT 1291
|||||
QY 1021 TGTGAGGCTCTTCGCAAGAGAAAGTGTCAAAACATCTCCACGGGCAAGCGCCTCC 1080
Db 1292 TGTGAGGCTCTTCGCAAGAGAAAGTGTCAAAACATCTCCACGGGCAAGCGCCTCC 1351
|||||
QY 1081 GTGAGCGGATCCCAAGTTACCCAGTAGCAAGGAGAGCCACTCACTGCTCCGAGTGC 1140
Db 1352 GTGAGCGGATCCCAAGTTACCCAGTAGCAAGGAGAGCCACTCACTGCTCCGAGTGC 1411
|||||
QY 1141 GGCAAAGCTTTCAAGAACCTACACAGCTGCTTTCGACTTCAGGCTCCACAAGAGAGAC 1200
Db 1412 GGCAAAGCTTTCAAGAACCTACACAGCTGCTTTCGACTTCAGGCTCCACAAGAGAGAC 1471
|||||
QY 1201 CGGAGGCGCGCGGAGTCCGCCAACCATGCTGTGGAAGGAGAGCGAGCGGGGAGCTGT 1260
Db 1472 CGGAGGCGCGCGGAGTCCGCCAACCATGCTGTGGAAGGAGAGCGAGCGGGGAGCTGT 1531
|||||
QY 1261 TCTCTGACCTCGCGGCGCTCTGGATCAAAATGAGAGCGGTGGATCGAGGGGAGGCTGT 1320
Db 1532 TCTCTGACCTCGCGGCGCTCTGGATCAAAATGAGAGCGGTGGATCGAGGGGAGGCTGT 1591
|||||
QY 1321 TCTGAAGACGATCTGAGGATGGGCTTCCCAAGGAATCCCATCTCGATAAAAAATCATGAT 1380
Db 1592 TCTGAAGACGATCTGAGGATGGGCTTCCCAAGGAATCCCATCTCGATAAAAAATCATGAT 1651
|||||
QY 1381 GGAGAAAAATTAACATCTTACATCTTCAAGAGAGTGTAGTATTGTGGAAGTTTTTC 1440
Db 1652 GGAGAAAAATTAACATCTTACATCTTCAAGAGAGTGTAGTATTGTGGAAGTTTTTC 1711
|||||
QY 1441 CGTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1500
Db 1712 CGTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1771
|||||
QY 1501 TGTGAATTTTGTGAATGCTGCGAGCGCAGAGACATCTCTGAGGTATCACTTGGAGAGA 1560
Db 1772 TGTGAATTTTGTGAATGCTGCGAGCGCAGAGACATCTCTGAGGTATCACTTGGAGAGA 1831
|||||

| | | | | | | | | | |
|---|------|--|------|--|--|--|--|--|--|
| Query Match 95.4%; Score 3040.6; DB 4; Length 5632; | | | | | | | | | |
| Best Local Similarity 95.8%; Pred. No. 0; | | | | | | | | | |
| Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1; | | | | | | | | | |
| QY | 1 | ATGCAATCGAAGTGCAGAGGAAACATGCCAACTCAATCCCTTTAAATGTATCATGGATGGG | 60 | | | | | | |
| DB | 272 | ATGCAATCGAAGTGCAGAGGAAACATGCCAACTCAATCCCTTTAAATGTATCATGGATGGG | 331 | | | | | | |
| QY | 61 | CCAGAAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG | 120 | | | | | | |
| DB | 332 | CCAGAAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG | 391 | | | | | | |
| QY | 121 | AAAGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCTCATCCAAATCGAG | 180 | | | | | | |
| DB | 392 | AAAGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCTCATCCAAATCGAG | 451 | | | | | | |
| QY | 181 | GGGTATATGCCCTTGGATTTGCAATGTTCTGACGACGACCTTCACACATTCAGAAGACCTT | 240 | | | | | | |
| DB | 452 | GGGTATATGCCCTTGGATTTGCAATGTTCTGACGACGACCTTCACACATTCAGAAGACCTT | 511 | | | | | | |
| QY | 241 | AATAAACATGTCTTAATCAACACCGGCTTACCTCTGTGAAACAGACAGTCTTTCGGGTT | 300 | | | | | | |
| DB | 512 | AATAAACATGTCTTAATGCAACACCGGCTTACCTCTGTGAAACAGCAGTCTTTCGGGTT | 571 | | | | | | |
| QY | 301 | GAAGCAGAGTATCTCAGTCCGCTTGATTAAGTCAAGTGGGACAGAACTTCCCAAGGAA | 360 | | | | | | |
| DB | 572 | GAAGCAGAGTATCTCAGTCCGCTTGATTAAGTCAAGTGGGACAGAACTTCCCAAGGAA | 631 | | | | | | |
| QY | 361 | AAGAAATGCAAGGAAATGAATTTAGTGTGAGGTATGTGGGACACATTTAGAGTCGCT | 420 | | | | | | |
| DB | 632 | AAGAAATGCAAGGAAATGAATTTAGTGTGAGGTATGTGGGACACATTTAGAGTCGCT | 691 | | | | | | |
| QY | 421 | TTTGATGTTGAGATCCACATGAGAAACACAAAGATTTCTTCACTTACGGGTGAACATG | 480 | | | | | | |
| DB | 692 | TTTGATGTTGAGATCCACATGAGAAACACAAAGATTTCTTCACTTACGGGTGAACATG | 751 | | | | | | |
| QY | 481 | TGCGGAAGAAGMTSRRSSAGGCTTGGTTTCTTAAAAATCAATCGGACACATAATGGC | 540 | | | | | | |
| DB | 752 | TGCGGAAGAAGATTCAAGGAGGCTTGGTTTCTTAAAAATCAATCGGACACATAATGGC | 811 | | | | | | |
| QY | 541 | AAATCGGGGCGCAGAGCAAACTGCAGCAAGGCTTGGAGGTAGTCCAGCAACGATCAAC | 600 | | | | | | |
| DB | 812 | AAATCGGGGCGCAGAGCAAACTGCAGCAAGGCTTGGAGGTAGTCCAGCAACGATCAAC | 871 | | | | | | |
| QY | 601 | GAGGTGCTCCAGGTGCAACGCGCGCAGAGCATCTCTCTCTTACAAATCTGCATGGTT | 660 | | | | | | |
| DB | 872 | GAGGTGCTCCAGGTGCAACGCGCGCAGAGCATCTCTCTCTTACAAATCTGCATGGTT | 931 | | | | | | |
| QY | 661 | TGTGGCTTCTTATTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA | 720 | | | | | | |
| DB | 932 | TGTGGCTTCTTATTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA | 991 | | | | | | |
| QY | 721 | AAAACTGCTTTTCGGTACCAGCAGCGCAGACAGACTCTCCACAGGAGGAAATGCCGTCC | 780 | | | | | | |
| DB | 992 | AAAACTGCTTTTCGGTACCAGCAGCGCAGACAGACTCTCCACAGGAGGAAATGCCGTCC | 1051 | | | | | | |
| QY | 781 | TCGAGGAGGACTTCTCGCAGTTGTTCAACTTGAGACCAAAATCTCACCTGAAACCGGG | 840 | | | | | | |
| DB | 1052 | TCGAGGAGGACTTCTCGCAGTTGTTCAACTTGAGACCAAAATCTCACCTGAAACCGGG | 1111 | | | | | | |
| QY | 841 | AAGAAGCTGTGCAGATGCATCCCTCAGCTTCGATTCGTTTACCACCTTCCAGCTTGGCAG | 900 | | | | | | |
| DB | 1112 | AAGAAGCTGTGCAGATGCATCCCTCAGCTTCGATTCGTTTACCACTTCCAGCTTGGCAG | 1171 | | | | | | |
| QY | 901 | CTGGCTACCAAGGAAAAAGTTGCCATTTGCCAAGAGTGAAGGATTCGGGGCAGAAGGG | 960 | | | | | | |
| DB | 1172 | CTGGCTACCAAGGAAAAAGTTGCCATTTGCCAAGAGTGAAGGATTCGGGGCAGAAGGG | 1231 | | | | | | |
| QY | 961 | AGCACCGCAACACGAGTTTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATAAGGGCAGT | 1020 | | | | | | |
| DB | 1232 | AGCACCGCAACACGAGTTTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATAAGGGCAGT | 1291 | | | | | | |
| QY | 1021 | TGTGAGGCTCTTCGCAAGAGAAAGAGAGTGCAAAACACTCCCAAGGAGGCGCCCTCC | 1080 | | | | | | |
| DB | 1292 | TGTGAGGCTCTTCGCAAGAGAAAGAGAGTGCAAAACACTCCCAAGGAGGCGCCCTCC | 1351 | | | | | | |
| QY | 1081 | GTGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGTCTCCAGTGC | 1140 | | | | | | |
| DB | 1352 | GTGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGTCTCCAGTGC | 1411 | | | | | | |
| QY | 1141 | GGCAAGCTTTTCAGAACCTTACCAAGCTGTGTTCACCTCAGGGTCCACAAGAGGAC | 1200 | | | | | | |
| DB | 1412 | GGCAAGCTTTTCAGAACCTTACCAAGCTGTGTTCACCTCAGGGTCCACAAGAGGAC | 1471 | | | | | | |
| QY | 1201 | CGGAGGCGCGGCGGAGTCCGCCACCATGTCTGTGACGGGAGGACGCCGGGAGCTGT | 1260 | | | | | | |
| DB | 1472 | CGGAGGCGCGGCGGAGTCCGCCACCATGTCTGTGACGGGAGGACGCCGGGAGCTGT | 1531 | | | | | | |
| QY | 1261 | TCTCTGACCTCGCGGCCCTCTGGATGAAATGGAGCGGTGGATCGAGGGGAGGTGT | 1320 | | | | | | |
| DB | 1532 | TCTCTGACCTCGCGGCCCTCTGGATGAAATGGAGCGGTGGATCGAGGGGAGGTGT | 1591 | | | | | | |
| QY | 1321 | TCTGAAGCGGATCTGAGGATGGGCTTCCGAAAGGAATCCATCTGGATTAATAATGATGAT | 1380 | | | | | | |
| DB | 1592 | TCTGAAGCGGATCTGAGGATGGGCTTCCGAAAGGAATCCATCTGGATTAATAATGATGAT | 1651 | | | | | | |
| QY | 1381 | GGAGGAAAAATAAACAATCTTACATCTTCAAGAGAGTGTAGTTATTTGGGAAAGTTTTTC | 1440 | | | | | | |
| DB | 1652 | GGAGGAAAAATAAACAATCTTACATCTTCAAGAGAGTGTAGTTATTTGGGAAAGTTTTTC | 1711 | | | | | | |
| QY | 1441 | CGTTCAAAATTTACCTCAATATTTATCTCAGAACGCATACAGGTGAAAAACCATACAAA | 1500 | | | | | | |
| DB | 1712 | CGTTCAAAATTTACCTCAATATTTATCTCAGAACGCATACAGGTGAAAAACCATACAAA | 1771 | | | | | | |
| QY | 1501 | TGTGAATTTTGTGAATATGTGCGAGCCAGAGACATCTCTGAGGTATCACTTGGAGAGA | 1560 | | | | | | |
| DB | 1772 | TGTGAATTTTGTGAATATGTGCGAGCCAGAGACATCTCTGAGGTATCACTTGGAGAGA | 1831 | | | | | | |
| QY | 1561 | CATCAAGGAAAAAACAACCGATGTTGCTGCTGAAAGTCAAGAACGATGGTAAAAATCAG | 1620 | | | | | | |
| DB | 1832 | CATCAAGGAAAAAACAACCGATGTTGCTGCTGAAAGTCAAGAACGATGGTAAAAATCAG | 1891 | | | | | | |
| QY | 1621 | GACACTGAAGTGCATTAATAACCGCTGACAGTGGCGAAACCAAAAATTTGAAAAATTT | 1680 | | | | | | |
| DB | 1892 | GACACTGAAGTGCATTAATAACCGCTGACAGTGGCGAAACCAAAAATTTGAAAAATTT | 1951 | | | | | | |
| QY | 1681 | TTTGATGTTGCCAAGATGTTTACAGGAGTCCACCTGCAAAAGCAGCTTAAAGSAGATGCT | 1740 | | | | | | |
| DB | 1952 | TTTGATGTTGCCAAGATGTTTACAGGAGTCCACCTGCAAAAGCAGCTTAAAGSAGATGCT | 2011 | | | | | | |
| QY | 1741 | TCTGTTTTTCAAGATGTTCTGGGAGCGCTGTCTCTCTCACCAGCACACAAGATACCTAG | 1800 | | | | | | |
| DB | 2012 | TCTGTTTTTCAAGATGTTCTGGGAGCGCTGTCTCTCTCACCAGCACACAAGATACCTAG | 2071 | | | | | | |
| QY | 1801 | GATTTCCATAAAATGCAAGTGCATGCTGATAAGTGAATTAATAAATTTGAAAAATTT | 1860 | | | | | | |
| DB | 2072 | GATTTCCATAAAATGCAAGTGCATGCTGATAAGTGAATTAATAAATTTGAAAAATTT | 2131 | | | | | | |
| QY | 1861 | GCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAACCTCATC | 1920 | | | | | | |
| DB | 2132 | GCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAACCTCATC | 2191 | | | | | | |
| QY | 1921 | TGTAGAACCAAGGCGGATGTTTACTCTCTCGGATGCGAGTACCAACCATACCTTGA | 1980 | | | | | | |
| DB | 2192 | TGTAGAACCAAGGCGGATGTTTACTCTCTCGGATGCGAGTACCAACCATACCTTGA | 2251 | | | | | | |
| QY | 1981 | GTTAGCCCCCAAGAGAGCAAAACGAGACCGCAGCTGACTGCAGATACAGGCAAGTGTG | 2040 | | | | | | |
| DB | 2252 | GTTAGCCCCCAAGAGAGCAAAACGAGACCGCAGCTGACTGCAGATACAGGCAAGTGTG | 2311 | | | | | | |
| QY | 2041 | GATTTGTCAAGAAAAACCTTTAAATTTATTCGTTGGGGGCTCTTTCACAAATTTGCCGCAAT | 2100 | | | | | | |
| DB | 2312 | GATTTGTCAAGAAAAACCTTTAAATTTATTCGTTGGGGGCTCTTTCACAAATTTGCCGCAAT | 2371 | | | | | | |
| QY | 2101 | TCTTTGAGTAAAGTTTGAATTCACCAAGTATACCTGTCCATTTTGTACCTTCAAGACATTT | 2160 | | | | | | |

Db 2372 TCTTTGAGTAAAGTTTGAATCCCAAGTATCACCTGTCTCATTTTGTACCTTCAAGACATTT 2431
Qy 2161 TATCCAGAGTTTAAATGATGACACAGAGACTGGAGCATAAATACAAATCCCTGAGCTTCAT 2220
Db 2432 TATCCAGAGTTTAAATGATGACACAGAGACTGGAGCATAAATACAAATCCCTGAGCTTCAT 2491
Qy 2221 AAAAAGCTGCAAAACAGTCTTGTGTAGAGTTCGACGTACCGGATGCCGCCAGCGTTG 2280
Db 2492 AAAAAGCTGCAAAACAGTCTTGTGTAGAGTTCGACGTACCGGATGCCGCCAGCGTTG 2551
Qy 2281 CTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2340
Db 2552 CTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2611
Qy 2341 CCGGGCAGTCCAAATCCCTGCCATCTCGAGGGGAGCAGAGCCCTCTGGGCCAGGC 2400
Db 2612 CCGGGCAGTCCAAATCCCTGCCATCTCGAGGGGAGCAGAGCCCTCTGGGCCAGGC 2671
Qy 2401 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAGTAACTGAAGTCC 2460
Db 2672 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAGTAACTGAAGTCC 2731
Qy 2461 CACAGACACAGCAGAAATGTGGGGTCAAGGGGCGGCCACAGGCCAACGAATCTGAG 2520
Db 2732 CACAGACACAGCAGAAATGTGGGGTCAAGGGGCGGCCACAGGCCAACGAATCTGAG 2791
Qy 2521 ATGTTTCTTAAACCAAGTGTTCCTCCCTCAGCGGATAGACAAAAGACCCGACACAAA 2580
Db 2792 ATGTTTCTTAAACCAAGTGTTCCTCCCTCAGCGGATAGACAAAAGACCCGACACAAA 2851
Qy 2581 TTGAAACCTCTTCAGTAGTCTCTTCTCAGGCCACCTCGGACAGTAACTCAATGTT 2640
Db 2852 TTGAAACCTCTTCAGTAGTCTCTTCTCAGGCCACCTCGGACAGTAACTCAATGTT 2911
Qy 2641 TCCATCGACTACCCGCCCAAGAACGACAGCCCGTGGGCACCTCCGGGAAGAGACTATTTTC 2700
Db 2912 TCCATCGACTACCCGCCCAAGAACGACAGCCCGTGGGCACCTCCGGGAAGAGACTATTTTC 2971
Qy 2701 TGTAACTCGAGTCCAGCAATACTGACAGAGAAATTTGTGTAGCCCTTCCAAAAGACTG 2760
Db 2972 TGTAACTCGAGTCCAGCAATACTGACAGAGAAATTTGTGTAGCCCTTCCAAAAGACTG 3031
Qy 2761 AAGTCCAGCTGTTGCTTGTGACCTTCAGCCCGGGGCCAAATTCAGAAAGAGGCTAT 2820
Db 3032 AAGTCCAGCTGTTGCTTGTGACCTTCAGCCCGGGGCCAAATTCAGAAAGAGGCTAT 3091
Qy 2821 GACCTTCCCAAGTACCATATGTTGACAGGCAATCACTGTTTACCAGGACTGTGTG 2880
Db 3092 GACCTTCCCAAGTACCATATGTTGACAGGCAATCACTGTTTACCAGGACTGTGTG 3151
Qy 2881 TATCCGTGCGAGGCGCTGCTCCCAACCAAGTTCTTGAGCTCCAGCGAGTGCATTTCT 2940
Db 3152 TATCCGTGCGAGGCGCTGCTCCCAACCAAGTTCTTGAGCTCCAGCGAGTGCATTTCT 3211
Qy 2941 CCAATGTGCTGACTGTTTCAGAGGCCCTATGTTGCTCCGGGCCACTTTACACTTGTGTG 3000
Db 3212 CCAATGTGCTGACTGTTTCAGAGGCCCTATGTTGCTCCGGGCCACTTTACACTTGTGTG 3271
Qy 3001 CTGCTGTGTAGTCCAGCATCCAGCTCGACGTTAG----- 3034
Db 3272 CTGCTGTGTAGTCCAGCATCCAGCTCGACGTTAG----- 3331
Qy 3035 ----- 3034
Db 3332 CACTTATCTAACAGCATGGCAACAAAGAGAACTATGAAATTTTATTTGGGAATGCAT 3391
Qy 3035 ----- 3047
Db 3392 TATCGACCAATGACAAAAAAACTTGAATTCATTAATTAGGGGAAAAAGGCTCTTGGTGG 3451
Qy 3048 ATGTCAGTGTCTTACCTCCCATGAAATTAATTTTACTTCTTGTGAGAGCGAATGTT 3107

Db 3452 ATGTCAGTGTCTTACTCCCATGAAATTAATTTTACTTCTCATCTTTGAGAGCAATGGT 3511
Qy 3108 GAAAGCTACTGAAATAAGCTGTGATTTCTACTGTACATAAAACATATCAGGAATCTGCAAG 3167
Db 3512 GAAAGCTACTGAAATAAGCTGTGATTTCTACTGTACATAAAACATATGAGGAATCTGCAAG 3571
Qy 3168 GAACACTACAGTTGTGTAA 3186
Db 3572 GAACACTACAGTTGTGTAA 3590
RESULT 5
US-09-949-016-12604
; Sequence 12604, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12604
; LENGTH: 20022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12604
Query Match 49.0%; Score 1561.2; DB 4; Length 20022;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1482 AGGTGAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGAGACATCTCT 1541
Db 7816 AGGTGAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGAGACATCTCT 7875
Qy 1542 GAGGTATCATTGAGAGACATCAAGAGAAAAACCAACCATGTTGCTGCTGAAGTCAA 1601
Db 7876 GAGGTATCATTGAGAGACATCAAGAGAAAAACCAACCATGTTGCTGCTGAAGTCAA 7935
Qy 1602 GAACGATGGTAAAAATCAGGACACTGAAGTGCATTTAAACCGTGAAGTGCAGCAAC 1661
Db 7936 GAACGATGGTAAAAATCAGGACACTGAAGTGCATTTAAACCGTGAAGTGCAGCAAC 7995
Qy 1662 CAAAAATTTGAAAGATTTTGTGATGTTGCAAGATGTTTACAGGAGTCCACCTGCAAA 1721
Db 7996 CAAAAATTTGAAAGATTTTGTGATGTTGCAAGATGTTTACAGGAGTCCACCTGCAAA 8055
Qy 1722 GCAGCTTAAGGAGATGCTTCTGTTTTCAGAAATGTTTGGGCGAGCGTGTCTCTCACC 1781
Db 8056 GCAGCTTAAGGAGATGCTTCTGTTTTCAGAAATGTTTGGGCGAGCGTGTCTCTCACC 8115
Qy 1782 AGCACAAGAGATCTCAGGATTTCCATAAAATGCAAGTGCAGTGCATGATGATAAAGT 1841
Db 8116 AGCACAAGAGATCTCAGGATTTCCATAAAATGCAAGTGCAGTGCATGATAAAGT 8175
Qy 1842 GAATAAAAACCTACCTGCTTACCTGACCTGTTAAAAAGATCAGCAGTTGAAAC 1901
Db 8176 GAATAAAAACCTACCTGCTTACCTGACCTGTTAAAAAGATCAGCAGTTGAAAC 8235
Qy 1902 TCAGGCAATTAACCTCATCTCTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG 1961
Db 8236 TCAGGCAATTAACCTCATCTCTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG 8295
Qy 1962 TACCACCCATTAACCTTGAAGTTAGCCCCCAAGAGAAAGCAACCGGAGCCGAGCTGACTG 2021

Db 8296 TACCACCCATACCTTGAAGTTAGCCCAAGAGAAAGCAACCGAGACGCGAGCTGACTG 8355
QY 2022 CAGATACAGGCCAAGTGTGGATGTGTACGAAAAAAGCTTTAAATTTATCGTGGGGGCTCT 2081
Db 8356 CAGATACAGGCCAAGTGTGGATGTGTACGAAAAAAGCTTTAAATTTATCGTGGGGGCTCT 8415
QY 2082 TCACAAATGCGCGCAATTTCTTCAGTAAGATTTGATTCAGATATCACTGTCCTATT 2141
Db 8416 TCACAAATGCGCGCAATTTCTTCAGTAAGATTTGATTCAGATATCACTGTCCTATT 8475
QY 2142 TTGTACCTTCAAGACATTTATCCAGAAGTTTAAATGATGACACAGAGACTGGACATAA 2201
Db 8476 TTGTACCTTCAAGACATTTATCCAGAAGTTTAAATGATGACACAGAGACTGGACATAA 8535
QY 2202 ATACAAATCTGACGTTTCAATAAACTGTGGAACCAAGTCTTGTGTTAGAAATCGAGTAC 2261
Db 8536 ATACAAATCTGACGTTTCAATAAACTGTGGAACCAAGTCTTGTGTTAGAAATCGAGTAC 8595
QY 2262 CGGATGCGCGCAGCGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 2321
Db 8596 CGGATGCGCGCAGCGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 8655
QY 2322 CAAGCCCAAGTGTCTTCCCGCGCAGTCCAAATCCCTGCTGCAAGGGGGAAGCA 2381
Db 8656 CAAGCCCAAGTGTCTTCCCGCGCAGTCCAAATCCCTGCTGCAAGGGGGAAGCA 8715
QY 2382 GAGCCCTCTGCGCCAGGCAAGCCCTCTGACTTCAGGGATGAGACTTAGCATTGTTAGC 2441
Db 8716 GAGCCCTCTGCGCCAGGCAAGCCCTCTGACTTCAGGGATGAGACTTAGCATTGTTAGC 8775
QY 2442 CCCAAGTAACCTGAAGTCCACAGACACAGCAGAGATGTGGGGTCCAAAGGGGCGCCAC 2501
Db 8776 CCCAAGTAACCTGAAGTCCACAGACACAGCAGAGATGTGGGGTCCAAAGGGGCGCCAC 8835
QY 2502 CAGGCAACAGCAATCTGAGATGTTTCTTAAACAGAGTGTTCCTCCCTGACCGGTAAGAC 2561
Db 8836 CAGGCAACAGCAATCTGAGATGTTTCTTAAACAGAGTGTTCCTCCCTGACCGGTAAGAC 8895
QY 2562 AAAAGACCCGAGACAAAATGAAACCTTCTCAGTAGTCTCTTCCAGCCCAACCCCTCGG 2621
Db 8896 AAAAGACCCGAGACAAAATGAAACCTTCTCAGTAGTCTCTTCCAGCCCAACCCCTCGG 8955
QY 2622 CAGCAGTAACATCAATGTTCCATCGACTACCCCGCAAGACGACGCGGTGGGCAAC 2681
Db 8956 CAGCAGTAACATCAATGTTCCATCGACTACCCCGCAAGACGACGCGGTGGGCAAC 9015
QY 2682 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATCTGCGAGCAGAAATTTGGTGA 2741
Db 9016 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATCTGCGAGCAGAAATTTGGTGA 9075
QY 2742 GCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCCCTTGACGTTGACAGCCCGGGGC 2801
Db 9076 GCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCCCTTGACGTTGACAGCCCGGGGC 9135
QY 2802 CAAATACAGAGAGGCTATGACCTTCCCAAGTACCATATGTTGAGAGCATCAATCACT 2861
Db 9136 CAAATACAGAGAGGCTATGACCTTCCCAAGTACCATATGTTGAGAGCATCAATCACT 9195
QY 2862 GTTACCGCAGGACTGTGTATCCGTGCGAGCGCTGCTCCCAACCAAGGTTCTTGAG 2921
Db 9196 GTTACCGCAGGACTGTGTATCCGTGCGAGCGCTGCTCCCAACCAAGGTTCTTGAG 9255
QY 2922 CTCCAGCAGGCTGCAATTTCCAAATGTGCTGACTGTTCCAGAGCCCTATGGTGGTCCGG 2981
Db 9256 CTCCAGCAGGCTGCAATTTCCAAATGTGCTGACTGTTCCAGAGCCCTATGGTGGTCCGG 9315
QY 2982 GCCACTTTACACTTGTGTGCTGTGTTAGTCCAGCATCCAGCTCGAGTTAGAGGTCT 3041
Db 9316 GCCACTTTACACTTGTGTGCTGTGTTAGTCCAGCATCCAGCTCGAGTTAGAGGTAT 9375
QY 3042 TGGTGGATGTCAGT 3055

Db 9376 TGCATGAGGGGCGT 9389
RESULT 6
US-09-949-016-16004
; Sequence 16004, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16004
; LENGTH: 20023
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16004
Query Match 49.0%; Score 1561.2; DB 4; Length 20023;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGACATCTCT 1541
Db 7816 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGACATCTCT 7875
QY 1542 GAGGTATCATTGAGAGACATCAACAGGAAAAAACAACCGATGTTGCTGCTGAAGTCAA 1601
Db 7876 GAGGTATCATTGAGAGACATCAACAGGAAAAAACAACCGATGTTGCTGCTGAAGTCAA 7935
QY 1602 GAACGATGTTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCAAC 1661
Db 7936 GAACGATGTTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCAAC 7995
QY 1662 CAAAAATTTGAAAAAGATTTTGTGATGTGCAAAAGATGTTTACAGGCACTGCACTGCAAA 1721
Db 7996 CAAAAATTTGAAAAAGATTTTGTGATGTGCAAAAGATGTTTACAGGCACTGCACTGCAAA 8055
QY 1722 GCAGCTTAAGAGATGCTTCTGTTTTCAGAAATGTTCTGGGCAAGCTGCTCTCACC 1781
Db 8056 GCAGCTTAAGAGATGCTTCTGTTTTCAGAAATGTTCTGGGCAAGCTGCTCTCACC 8115
QY 1782 AGCACACAAAGATACTCAGGATTTCCATAAAATGCACTGATGACAGTGTGATAAAGT 1841
Db 8116 AGCACACAAAGATACTCAGGATTTCCATAAAATGCACTGATGACAGTGTGATAAAGT 8175
QY 1842 GAATAAAAACCTTACCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 1901
Db 8176 GAATAAAAACCTTACCCTGCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAC 8235
QY 1902 TCAGGCAAAATCACTCATCTGTAGAACCAAGGCGGATGTTACTCCTCTCCGGATGGAG 1961
Db 8236 TCAGGCAAAATCACTCATCTGTAGAACCAAGGCGGATGTTACTCCTCTCCGGATGGAG 8295
QY 1962 TACCACCCATAACCTTCAAGTTAGCCCCCAAGAGACAAACGAGACCGCAGCTGACTG 2021
Db 8296 TACCACCCATAACCTTCAAGTTAGCCCCCAAGAGACAAACGAGACCGCAGCTGACTG 8355
QY 2022 CAGATACAGGCCAAGTGTGGATTTGTCACGAAAAACCTTTTAAATTTATCCGTTGGGGCTCT 2081
Db 8356 CAGATACAGGCCAAGTGTGGATTTGTCACGAAAAACCTTTTAAATTTATCCGTTGGGGCTCT 8415
QY 2082 TCACAAATGCGCGCAATTTCTTTTGTAGTAAAGTTTGTGATTCCAAGTATCACCTGTCCATT 2141

Db 8416 TCACAAATGCCCGCAATTTCTTTGAGTAAAGATTGATTCACAAATGATCACCTGTCCATT 8475
Qy 2142 TTGTACCTTCAAGACATTTTATCCAGAGTTTATATGATGACACAGAGACTGGAGCATAA 2201
Db 8476 TTGTACCTTCAAGACATTTTATCCAGAGTTTATATGATGACACAGAGACTGGAGCATAA 8535
Qy 2202 ATACAATCCTGACGTTTCATAAAAACTGTCGAAAACAAAGTCCCTTGCTTGAAGTGCACGTAC 2261
Db 8536 ATACAATCCTGACGTTTCATAAAAACTGTCGAAAACAAAGTCCCTTGCTTGAAGTGCACGTAC 8595
Qy 2262 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCTCCCTCTCTCTAGTTTCTGTAAACC 2321
Db 8596 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCTCCCTCTCTCTAGTTTCTGTAAACC 8655
Qy 2322 CAAGCCCAAGTCTGCTTTCCCGCGCGCAGTCCAAATCCCTGCAATCTGCGAAGGGGAAGCA 2381
Db 8656 CAAGCCCAAGTCTGCTTTCCCGCGCGCAGTCCAAATCCCTGCAATCTGCGAAGGGGAAGCA 8715
Qy 2382 GAGCCCTCCTGGCGCAGCAAGGCCCTCTGACTTTCCAGGATAGACTCTAGCACTTTAGC 2441
Db 8716 GAGCCCTCCTGGCGCAGCAAGGCCCTCTGACTTTCCAGGATAGACTCTAGCACTTTAGC 8775
Qy 2442 CCCAAGTAACTGAGTCCCAAGACACAGCAAGATGTGGGGTCCAAAGGGGCGGCCAC 2501
Db 8776 CCCAAGTAACTGAGTCCCAAGACACAGCAAGATGTGGGGTCCAAAGGGGCGGCCAC 8835
Qy 2502 CAGCACAAGCAATCTGAGATGTTTCTTAAACAGAGTGTTCCTCCCTGACCGGATAGAC 2561
Db 8836 CAGGCAACAGCAATCTGAGATGTTTCTTAAACAGAGTGTTCCTCCCTGACCGGATAGAC 8895
Qy 2562 AAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCACCCCTCGG 2621
Db 8896 AAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGCTCCTTCTCAGCCACCCCTCGG 8955
Qy 2622 CAGCAGTAACTAATGTTTCATCGACTACCCCGCCAAAGAACGACACGCCGTGGGCCACC 2681
Db 8956 CAGCAGTAACTAATGTTTCATCGACTACCCCGCCAAAGAACGACACGCCGTGGGCCACC 9015
Qy 2682 TCCGGGAGAGACTATTTCTGTAATCGGAGTGCAGCAATCTGCGAGCAATTTGGTGA 2741
Db 9016 TCCGGGAGAGACTATTTCTGTAATCGGAGTGCAGCAATCTGCGAGCAATTTGGTGA 9075
Qy 2742 GCCCCTTCCAAAAAGACTGAAAGTCAGCGGTGGTTGCCCTTGACGTTGACCGCCCGGGGC 2801
Db 9076 GCCCCTTCCAAAAAGACTGAAAGTCAGCGGTGGTTGCCCTTGACGTTGACCGCCCGGGGC 9135
Qy 2802 CAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTTGAGGAGCATCATCACT 2861
Db 9136 CAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTTGAGGAGCATCATCACT 9195
Qy 2862 GTTACCGCAGACTGTGTATCCGTGCGAGCGCTGCCCTCCCAACCAAGGTTCCCTGAG 2921
Db 9196 GTTACCGCAGACTGTGTATCCGTGCGAGCGCTGCCCTCCCAACCAAGGTTCCCTGAG 9255
Qy 2922 CTCAGCAGGTCGATTTCCAAATGTGCTGACTGTTCAGAGCCCTATGTTGGCTCCGG 2981
Db 9256 CTCAGCAGGTCGATTTCCAAATGTGCTGACTGTTCAGAGCCCTATGTTGGCTCCGG 9315
Qy 2982 GCCACTTTACACTTGTGTGCTGCTGGTAGTCCAGCATCCAGCTCGAGTTAGAAAGTCT 3041
Db 9316 GCCACTTTACACTTGTGTGCTGCTGGTAGTCCAGCATCCAGCTCGAGTTAGAAAGTAT 9375
Qy 3042 TGGTGGATGTCAGT 3055
Db 9376 TGCATGAGGGGGCGT 9389

RESULT 7
US-08-892-695-9
; Sequence 9, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:

; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Kommen, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892.695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 10365
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Genomic
; OTHER INFORMATION: Sequence encoding ZABC1
; Patent No. 6808878
US-08-892-695-9

Query Match 49.0%; Score 1559.6; DB 4; Length 10365;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTCAGCCAGAGACATCTCT 1541
Db 8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTCAGCCAGAGACATCTCT 8345
Qy 1542 GAGTATCACTTGTGAGAGACATCAAGAGAAAAACAAACCGATGTTGCTGTAAGTCAA 1601
Db 8346 GAGTATCACTTGTGAGAGACATCAAGAGAAAAACAAACCGATGTTGCTGTAAGTCAA 8405
Qy 1602 GAACGATGTTAAATCAGGACACTGAAGATGCACTATTAAACCGTACAGTGGCGCAAC 1661
Db 8406 GAACGATGTTAAATCAGGACACTGAAGATGCACTATTAAACCGTACAGTGGCGCAAC 8465
Qy 1662 CAAAAATTTGAAAAAGATTTTTTGTGATGTTGTCGCAAGATGTTACAGGAGTCAACCTGCAAA 1721
Db 8466 CAAAAATTTGAAAAAGATTTTTTGTGATGTTGTCGCAAGATGTTACAGGAGTCAACCTGCAAA 8525
Qy 1722 GCAGTTAAGAGATGCTCTGTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACC 1781
Db 8526 GCAGTTAAGAGATGCTCTGTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACC 8585
Qy 1782 AGCACACAAAGATCTCAGGATTTCCATAAAATGCACTGATGACAGTGTGATAAAGT 1841
Db 8586 AGCACACAAAGATCTCAGGATTTCCATAAAATGCACTGATGACAGTGTGATAAAGT 8645
Qy 1842 GAATAAAAAACCTTACCCTGCTTACCTGGAACCTGTTTAAAAAAGAGATCAGCAGTTGAAAC 1901
Db 8646 GAATAAAAAACCTTACCCTGCTTACCTGGAACCTGTTTAAAAAAGAGATCAGCAGTTGAAAC 8705
Qy 1902 TCAGGCAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCCGATGGGAG 1961
Db 8706 TCAGGCAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCCGATGGGAG 8765
Qy 1962 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAGCAACCGAGACCGGAGCTGACTG 2021
Db 8766 TACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAGCAACCGAGACCGGAGCTGACTG 8825
Qy 2022 CAGATACAGGCCAAAGTGTGGATTGTTCAGGAAAAACCTTTTAAATTTTATCCGTGGGGCTCT 2081
Db 8826 CAGATACAGGCCAAAGTGTGGATTGTTCAGGAAAAACCTTTTAAATTTTATCCGTGGGGCTCT 8885
Qy 2082 TCACAAATGCCCCGCAATTTCTTTGAGTAAAGTTTGAATTCGAAGTATCACTGTCCATT 2141

Db 618 TGAAGCAGAGTATCTCAGTCCGCTGTGATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGA 677
Qy 360 AAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGTGGGAGAGATTTAGAGTCGC 419
Db 678 AAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGTGGGAGAGATTTAGAGTCGC 737
Qy 420 TTTTGTATGTTGAGATCCATGAGAACACAAAGATTTCTTTCATCTACCGGGTGAACAT 479
Db 738 TTTTGTATGTTGAGATCCATGAGAACACAAAGATTTCTTTCATCTACCGGGTGAACAT 797
Qy 480 GTGCGGAGAGAGMTTSSRSAGCTTGGTTCTTTAAATAATCACATGCGGACACATAATGG 539
Db 798 GTGCGGAGAGAGATTTCAAGAGCTTGGTTCTTTAAATAATCACATGCGGACRATATGG 857
Qy 540 CAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAGTGAGCAACAGATCAA 599
Db 858 CAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAGTGAGCAACAGATCAA 917
Qy 600 CGAGGTGCTCAGGTGCAACGCGCGGAGAGATCTCTCTCTTACAAATCTCATGTT 659
Db 918 CGAGGTGCTCAGGTGCAACGCGCGGAGAGATCTCTCTCTTACAAATCTCATGTT 977
Qy 660 TTGTGGCTTCTATTTCCAAATAAAGAAAGTCTAATTTGAGCAGCGCAAGGTGCAACCAA 719
Db 978 TTGTGGCTTCTATTTCCAAATAAAGAAAGTCTAATTTGAGCAGCGCAAGGTGCAACCAA 1037
Qy 720 AAAAAGCTGCTTTCGATACAGCAGCGGCGAGAGACTCTCCACAAAGGAGAAATGCCGTC 779
Db 1038 AAAAAGCTGCTTTCGATACAGCAGCGGCGAGAGACTCTCCACAAAGGAGAAATGCCGTC 1097
Qy 780 CTCGAGGAGGACTTCTGCAAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGG 839
Db 1098 CTCGAGGAGGACTTCTGCAAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGG 1157
Qy 840 GAAGAGCTGTCAGATCATCCCTCAGCTCGATCCGTTTACCACTTCCAGGCTTGCA 899
Db 1158 GAAGAGCTGTCAGATCATCCCTCAGCTCGATCCGTTTACCACTTCCAGGCTTGCA 1217
Qy 900 GCTGCTACCAAGGAAAGTTGCCATTTGCAAGAGTGAAGAAATCGGGGCAAGAGG 959
Db 1218 KCTGCTACCAAGGAAAGTTGCCATTTGCAAGAGTGAAGAAATCGGGGCAAGAGG 1277
Qy 960 GAGCACCAGCAACGACGATTCAGTTCCGAGAGGAGCTTCGAGAAACAAATAAGGCGAG 1019
Db 1278 GAGCACCAGCAACGACGATTCAGTTCCGAGAGGAGCTTCGAGAAACAAATAAGGCA 1337
Qy 1020 TTGTGAGGCTTCTGCAAGAGAAAGAGAGTGCACACTCCACGCGAGCGCCCTC 1079
Db 1338 TTGTGAGGCTTCTGCAAGAGAAAGAGAGTGCACACTCCACGCGAGCGCCCTC 1397
Qy 1080 CGTGACCGGATCCCAAGTTTACCCAGTAGCAAGGAGAGCCACTCACTGCTCCGAGTG 1139
Db 1398 CGTGACCGGATCCCAAGTTTACCCAGTAGCAAGGAGAGCCACTCACTGCTCCGAGTG 1457
Qy 1140 CGGCAAGCTTTTCAAGACTTACCACGCTGCTTTCGACTCCAGGGTCC 1189
Db 1458 CGGCAAGCTTTTCAAGACTTACCACGCTGCTTTCGACTCCAGGGTCC 1507

RESULT 9

US-08-892-695-3
; Sequence 3, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A

; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; FEATURE: modified_base
; NAME/KEY: (447)
; LOCATION: (447)
; OTHER INFORMATION: N is A, G, T, or U
US-08-892-695-3

Query Match 34.7%; Score 1106.2; DB 4; Length 1507;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1150; Conservative 9; Mismatches 28; Indels 3; Gaps 3;

Qy 1 ATGCAATCGAAAGTGACAGGAACATGCCAACTCAATCCCTCTTAAATGTATCATGGTGG 60
Db 320 ATGCAATCGAAAGTGACAGGAACATGCCAACTCAATCCCTCTTAAATGTATCATGGTGG 378
Qy 61 CCAGAGTGTGGCAGCTCTCTTGGCAGTCCGATGAGATGGAGATGCTTCTCAATG 120
Db 379 CCAAGAGTGTGGCAGCTCTCTTGGCAGTCCGATGAGATGGAGATGGA-GATGCTTCTCAATG 437
Qy 121 AAAGGGACCGCTGTGTTCCATTCGAGCTACACAAAGAAAAAATGTC-ATCCAAATCGA 179
Db 438 AAAGGGCCNCTGTGTTCAATTCGAGCTACAAAGAAAAAATGTCATCCGATCGA 497
Qy 180 GGGGTATATGCCCTTGGATTCATGTTCTGAGCCAGACCTTTCACATTCAGAAAGCCT 239
Db 498 GGGGAATATGCCCTTGGATTCATGTTCTGAGCCAGACCTTTCACATTCAGAAAGCCT 557
Qy 240 TAATAACATGTTCTTAATGCACACCGGCTTACCTCTGTGAACAGCAGTCTTCCGGT 299
Db 558 TAATAACATGTTCTTAATGCACACCGGCTTACCTCTGTGAACAGCAGTCTTCCGGT 617
Qy 300 TGAAGCAGAGTATCTCAGTCGCTTGAATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGA 359
Db 618 TGAAGCAGAGTATCTCAGTCGCTTGAATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGA 677
Qy 360 AAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATGTGGGAGACATTTAGAGTCGC 419
Db 678 AAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATGTGGGAGACATTTAGAGTCGC 737
Qy 420 TTTTGTATGTTGAGATCCATGAGAACACAAAGATTTCTTTCATCTACCGGGTGAACAT 479
Db 738 TTTTGTATGTTGAGATCCATGAGAACACAAAGATTTCTTTCATCTACCGGGTGAACAT 797
Qy 480 GTGCGGAGAGAGMTTSSRSAGCTTGGTTCTTTAAATAATCACATGCGGACACATAATGG 539
Db 798 GTGCGGAGAGAGATTCAGGAGGCTTGGTTCTTTAAATAATCACATGCGGACRATATGG 857
Qy 540 CAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAGTGAGCAACAGATCAA 599
Db 858 CAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAGTGAGCAACAGATCAA 917
Qy 600 CGAGGTGCTCAGGTGCAACGCGCGGAGAGATCTCTCTCTTACAAATCTCATGTT 659
Db 918 CGAGGTGCTCAGGTGCAACGCGCGGAGAGATCTCTCTCTTACAAATCTCATGTT 977
Qy 660 TTGTGGCTTCTATTTCCAAATAAAGAAAGTCTAATTTGAGCAGCGCAAGGTGCAACCAA 719
Db 978 TTGTGGCTTCTATTTCCAAATAAAGAAAGTCTAATTTGAGCAGCGCAAGGTGCAACCAA 1037

Db 481 GCCTGTGAGATGATCCCTCAGTCGATCCGTTTCCACCACTTCCAGGCTTGGCAGCTGGC 422
Qy 906 TACCAAGGAAAGTTGCTTCATTTCCCAAGAGTGAAGGAATCGGGCAAGAGGAGCAGC 965
Db 421 TACCAAGGAAAGTTGCTTCATTTCCCAAGAGTGAAGGAATCGGGCAAGAGGAGCAGC 362
Qy 966 CGACAACGACGATTCGAGTTCCGAGAGGAGCTTTGGAGAAACAAATAGGCGAGTTGTGC 1025
Db 361 CGACAACGACGATTCGAGTTCCGAGAGGAGCTTTGGAGAAACAAATAGGCGAGTTGTGC 302
Qy 1026 AGGCTCTCCGAGAGAAAGAGTGAACACATCTCCACGCGAGAGCGCCCTCGGTGGA 1085
Db 301 WGGCTCTCCGAGAGAAAGAGTGAACACATCTCCACGCGAGAGCGCCCTCGGTGGA 242
Qy 1086 CGCGGATCCCAAGTTACCCAGTACGAGGAGGAGCCCACTCAGTCTCCGAGTGGCGCAA 1145
Db 241 CGCGGATCCCAAGTTACCCAGTACGAGGAGGAGCCCACTCAGTCTCCGAGTGGCGCAA 182
Qy 1146 AGCTTTCAAGAACCTACCAACAGCTGGTCTTGCACCTCCAGGGTCCACAAGAGGACCGGAG 1205
Db 181 AGCTTTCAAGAACCTACCAACAGCTGGTCTTGCACCTCCAGGGTCCACAAGAGGACCGGAG 122
Qy 1206 GCGCGGCGGAGTGCAGTCTGTGTGAAGGAGGAGCGGGGAGCGTGTCTCC 1265
Db 121 GCGCGGCGGAGTGCAGTCTGTGTGAAGGAGGAGCGGGGAGCGTGTCTCC 62
Qy 1266 TGACCTCGCCGCCCTCTGGATGAAATGGAGCGTGGATCGAGGGAGGTTGTTCTGA 1325
Db 61 TGACCTCGCCGCCCTCTGGATGAAATGGAGCGTGGATCGAGGGAGGTTGTTCTGA 2
Qy 1326 A 1326
Db 1 A 1

RESULT 12

US-09-949-016-34699/c
; Sequence 34699, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34699
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34699

Query Match 18.8%; Score 598.2; DB 4; Length 601;
Best Local Similarity 98.8%; Pred. No. 9.9e-175;
Matches 594; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 99 GATGAGGATGCTTGTCAATGAAAGGAGCAGCGCTGTTTCCATTCCGAGCTACACAAGA 158
Db 601 GATGAGGATGCTTGTCAATGAAAGGAGCAGCGCTGTTTCCATTCCGAGCTACACAAGA 542
Qy 159 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 218
Db 541 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 482
Qy 219 CTTACACATTGAGAAGACCTTAAATAAATGTCATGTTTAAATGCAACACCGGCTTACCCTCTG 278

Db 481 CTTTCAACAATTCAGAGACCTTAAATAAATGTCATGTTTAAATGCAACACCGGCTACCCCTCTG 422
Qy 279 TGAACCAAGCAGTCTTCCGGTGTGAAGCAGAGTATCTCAGTCCGCTTGCATAAAAGTCAAGT 338
Db 421 TGAACCAAGCAGTCTTCCGGTGTGAAGCAGAGTATCTCAGTCCGCTTGCATAAAAGTCAAGT 362
Qy 339 GCGAAACAGAACTCCCAAGGAAAGAAATTCGAAGAAATGAAATTTAGCTGTGAGGTATG 398
Db 361 GCGAAACAGAACTCCCAAGGAAAGAAATTCGAAGAAATGAAATTTAGCTGTGAGGTATG 302
Qy 399 TGGGCAACATTTAGAGTCGCTTTTGTATGTTGAGATCCATGAGAAACACAACAAGATTC 458
Db 301 YGGGCAACATTTAGAGTCGCTTTTGTATGTTGAGATCCATGAGAAACACAACAAGATTC 242
Qy 459 TTTTCACTTACGGGTGAACATGTCGGAAGAAAGTTSRRSAGCTTGTGTTCTTAAAAA 518
Db 241 TTTTCACTTACGGGTGAACATGTCGGAAGAAAGTTCAGGAGCTTGTGTTCTTAAAAA 182
Qy 519 TCACATCGGACACATATATGCAAAATCGGGGCGCAGAAAGCAAACTGCAGCAAGGCTTGA 578
Db 181 TCACATCGGACACATATATGCAAAATCGGGGCGCAGAAAGCAAACTGCAGCAAGGCTTGA 122
Qy 579 GAGTAGTCCAGCAACGATCAACAGAGTCTTCCAGGTGCAACGCGCCGAGAGCATCTCTC 638
Db 121 GAGTAGTCCAGCAACGATCAACAGAGTCTTCCAGGTGCAACGCGCCGAGAGCATCTCTC 62
Qy 639 TCCTTACAAATCTGCATGTTTGTGCTTCTTATTTCCAAATAAAGAAAGTCTTAATTGA 698
Db 61 TCCTTACAAATCTGCATGTTTGTGCTTCTTATTTCCAAATAAAGAAAGTCTTAATTGA 2
Qy 699 G 699
Db 1 G 1

RESULT 13

US-09-949-016-151879/c
; Sequence 151879, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151879
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-151879

Query Match 18.8%; Score 598.2; DB 4; Length 601;
Best Local Similarity 98.8%; Pred. No. 9.9e-175;
Matches 594; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 99 GATGAGGATGCTTGTCAATGAAAGGAGCAGCGCTGTTTCCATTCCGAGCTACACAAGA 158
Db 601 GATGAGGATGCTTGTCAATGAAAGGAGCAGCGCTGTTTCCATTCCGAGCTACACAAGA 542
Qy 159 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 218
Db 541 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 482


```
QY 219 CTTACACATTCAGAGACCTTAATAAATGCTTTAATGCAACACCGGCTTACCTCTG 278
Db 481 CTTACACATTCAGAGACCTTAATAAATGCTTTAATGCAACACCGGCTTACCTCTG 422
QY 279 TGAACAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 338
Db 421 TGAACAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 362
QY 339 GGAACAGACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGTCTGAGGTATG 398
Db 361 GGAACAGACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGTCTGAGGTATG 302
QY 399 TGGGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGAGAACACACAAAGATT 458
Db 301 YGGGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGAGAACACACAAAGATT 242
QY 459 TTTCACTTACCGGTGTAAATGTCGCGGAGAGMTTSSRRSAGCCTTGGTTTCTTAAAAA 518
Db 241 TTTCACTTACCGGTGTAAATGTCGCGGAGAGATTCAAGGAGCCTTGGTTTCTTAAAAA 182
QY 519 TCACATGGGACACATATGGAATCGGGGGCCAGAGCAAACTGCAAGCGGCTTGGG 578
Db 181 TCACATGGGACACATATGGAATCGGGGGCCAGAGCAAACTGCAAGCGGCTTGGG 122
QY 579 GAGTAGTCCAGCAACGATCAAGAGTCTCCAGGTGCAAGCGGCGGAGAGCATCTCCTC 638
Db 121 GAGTAGTCCAGCAACGATCAAGAGTCTCCAGGTGCAAGCGGCGGAGAGCATCTCCTC 62
QY 639 TCCCTTACAAATCTCATGTTTGGGCTTCTCTATTTCCAAATAAAGAAAGTCTAATTGA 698
Db 61 TCCCTTACAAATCTCATGTTTGGGCTTCTCTATTTCCAAATAAAGAAAGTCTAATTGA 2
QY 699 G 699
Db 1 G 1
```

RESULT 14

US-09-620-312D-61

; Sequence 61, Application US/09620312D

; Patent No. 6569662

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yundong

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: John Tillinghast

; APPLICANT: Dmanac, Radoje T.

; TITLE OF INVENTION: No. 6569662el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/09/620,312D

; CURRENT FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1105

; SOFTWARE: pt_FL_genes Version 1.0

; SEQ ID NO 61

; LENGTH: 2765

; TYPE: DNA

;

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (123)..(2291)
US-09-620-312D-61
```

Query Match 2.1%; Score 66.8; DB 4; Length 2765;

Best Local Similarity 59.4%; Pred. No. 3e-09; Mismatches 6; Indels 0; Gaps 0;

Matches 101; Conservative 6; Mismatches 63; Indels 0; Gaps 0;

QY 379 GAATTTAGCTGTGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCAC 438

Db 939 GAGTTCGCTGCCAAGTGTGCGGAGAGCTTTACACAGTCTTGTTTCTCAAGGGCCAC 998

QY 439 ATGAGACACACAAAGATTCTTTCATTACGGGTGTAAACATGTGCGGAGAGMTTSSRRS 498

Db 999 ATGCGTAGACAAAGGCTCTTCGATCATGCGTGTCCGGTGTGCGGCGCTGCTTCAAG 1058

QY 499 SAGCCTTGGTTCTTAAAAATCACATGCGGACACATATGCAATCGG 548

Db 1059 GAGCCCTGGTTCCTTAAAGAACACATGAAGGTGCACGCCAGCAAGCTGGG 1108

RESULT 15

US-09-949-016-4961

; Sequence 4961, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4961

; LENGTH: 2799

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-4961

Query Match 2.1%; Score 66.8; DB 4; Length 2799;

Best Local Similarity 59.4%; Pred. No. 3e-09; Mismatches 6; Indels 0; Gaps 0;

Matches 101; Conservative 6; Mismatches 63; Indels 0; Gaps 0;

QY 379 GAATTTAGCTGTGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCAC 438

Db 988 GAGTTCGCTGCCAAGTGTGCGGAGAGCTTTACACAGTCTTGTTTCTCAAGGGCCAC 1047

QY 439 ATGAGACACACAAAGATTCTTTCATTACGGGTGTAAACATGTGCGGAGAGMTTSSRRS 498

Db 1048 ATGCGTAGACAAAGGCTCTTCGATCATGCGTGTCCGGTGTGCGGCGCTGCTTCAAG 1107

QY 499 SAGCCTTGGTTCTTAAAAATCACATGCGGACACATATGCAATCGG 548

Db 1108 GAGCCCTGGTTCCTTAAAGAACACATGAAGGTGCACGCCAGCAAGCTGGG 1157

Search completed: July 3, 2005, 16:59:39

Job time : 509.445 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 16:40:17 ; Search time 1858.96 Seconds
(without alignments)
10749.639 Million cell updates/sec

Title: US-08-731-499-10
Perfect score: 3186
Sequence: 1 ATGCAATCGAAGTGACAGG.....GGACACTACAGTGTGTAA 3186

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-----------------------|-------------------|
| 1 | 3183.6 | 99.9 | 3186 | US-08-731-499-10 | Sequence 10, Appl |
| 2 | 3183.6 | 99.9 | 5252 | 13 US-10-087-192-1709 | Sequence 1709, Ap |
| 3 | 3040.6 | 95.4 | 5632 | 15 US-10-177-293-505 | Sequence 505, App |
| 4 | 1561.2 | 49.0 | 36022 | 13 US-10-087-192-1708 | Sequence 1708, Ap |
| 5 | 1559.6 | 49.0 | 10365 | 8 US-08-731-499-9 | Sequence 9, Appl |
| 6 | 1511.2 | 47.4 | 3016 | 13 US-10-087-192-1706 | Sequence 1706, Ap |
| 7 | 1106.2 | 34.7 | 1507 | 8 US-08-731-499-3 | Sequence 3, Appl |

| | | | | | | |
|------|-------|------|--------|----|---------------------|--------------------|
| 8 | 733 | 23.0 | 26345 | 13 | US-10-087-192-1705 | Sequence 1705, Ap |
| 9 | 444.4 | 13.9 | 469 | 13 | US-10-040-739-520 | Sequence 520, App |
| c 10 | 401.4 | 12.6 | 530 | 16 | US-10-029-386-9711 | Sequence 9711, Ap |
| c 11 | 251 | 7.9 | 251 | 16 | US-10-029-386-23411 | Sequence 23411, A |
| 12 | 164 | 5.1 | 267 | 9 | US-09-783-590-9057 | Sequence 9057, Ap |
| 13 | 102.4 | 3.2 | 6033 | 17 | US-10-172-118-1511 | Sequence 1511, Ap |
| 14 | 102.4 | 3.2 | 6033 | 18 | US-10-342-887-1511 | Sequence 1511, Ap |
| 15 | 102.4 | 3.2 | 8156 | 14 | US-10-074-475-93 | Sequence 93, Appl |
| 16 | 70 | 2.2 | 70 | 20 | US-10-758-307-110 | Sequence 110, Appl |
| 17 | 70 | 2.2 | 70 | 21 | US-10-852-797-372 | Sequence 372, App |
| c 18 | 66.8 | 2.1 | 1014 | 16 | US-10-029-386-20910 | Sequence 20910, A |
| c 19 | 66.8 | 2.1 | 1229 | 16 | US-10-029-386-20193 | Sequence 20193, A |
| c 20 | 66.8 | 2.1 | 2765 | 15 | US-10-037-270-61 | Sequence 61, Appl |
| 21 | 66.8 | 2.1 | 2765 | 17 | US-10-117-722-61 | Sequence 61, Appl |
| 22 | 66.8 | 2.1 | 3039 | 17 | US-10-172-118-1644 | Sequence 1644, Ap |
| 23 | 66.8 | 2.1 | 3039 | 18 | US-10-342-887-1644 | Sequence 1644, Ap |
| 24 | 66.8 | 2.1 | 3617 | 9 | US-09-764-847-1394 | Sequence 1394, Ap |
| 25 | 66.8 | 2.1 | 3617 | 9 | US-09-764-847-1396 | Sequence 1396, Ap |
| 26 | 66.8 | 2.1 | 3617 | 14 | US-10-092-154-1394 | Sequence 1394, Ap |
| 27 | 66.8 | 2.1 | 3617 | 14 | US-10-092-154-1396 | Sequence 1396, Ap |
| 28 | 63.4 | 2.0 | 5154 | 21 | US-10-956-157-1977 | Sequence 1977, Ap |
| 29 | 62.8 | 2.0 | 1033 | 16 | US-10-029-386-22959 | Sequence 22959, A |
| 30 | 62.8 | 2.0 | 2918 | 17 | US-10-104-047-835 | Sequence 835, App |
| 31 | 62.8 | 2.0 | 4599 | 10 | US-09-764-872-787 | Sequence 787, App |
| 32 | 62.8 | 2.0 | 4599 | 10 | US-09-764-872-790 | Sequence 790, App |
| 33 | 62.8 | 2.0 | 4603 | 10 | US-09-764-872-791 | Sequence 791, App |
| c 34 | 62.2 | 2.0 | 308 | 16 | US-10-029-386-15037 | Sequence 15037, A |
| c 35 | 62.2 | 2.0 | 592 | 16 | US-10-029-386-1335 | Sequence 1335, Ap |
| c 36 | 61.2 | 1.9 | 193691 | 20 | US-10-719-993-6768 | Sequence 6768, Ap |
| c 37 | 61.2 | 1.9 | 318760 | 20 | US-10-719-993-6765 | Sequence 6765, Ap |
| 38 | 60.6 | 1.9 | 3257 | 21 | US-10-956-157-1071 | Sequence 1071, Ap |
| 39 | 60 | 1.9 | 60 | 10 | US-09-308-975-12266 | Sequence 12266, A |
| 40 | 59.8 | 1.9 | 453 | 10 | US-09-918-995-5364 | Sequence 5364, App |
| 41 | 59.8 | 1.9 | 2703 | 17 | US-10-264-049-841 | Sequence 841, App |
| 42 | 59.8 | 1.9 | 2785 | 17 | US-10-120-988-105 | Sequence 105, App |
| 43 | 59.8 | 1.9 | 2901 | 9 | US-09-974-298-117 | Sequence 117, App |
| 44 | 59.8 | 1.9 | 3120 | 17 | US-10-094-749-337 | Sequence 337, App |
| 45 | 59.8 | 1.9 | 3646 | 10 | US-09-814-353-20524 | Sequence 20524, A |

ALIGNMENTS

RESULT 1

US-08-731-499-10
; Sequence 10, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3186 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..3186
; OTHER INFORMATION: /note= "ZABCl Open Reading Frame"
;
US-08-731-499-10

Query Match          99.9%; Score 3183.6; DB 8; Length 3186;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCAATCGAAGTACAGAGAAACATGCCAACTCAATCCCTTTAATGTACATGGATGGG 60
Db 1 ATGCAATCGAAGTACAGAGAAACATGCCAACTCAATCCCTTTAATGTACATGGATGGG 60

Qy 61 CCAGAAGTGATGGCAGCTCTCTGGCAGTCCGATCGGATGGAGATGGAGATGCTTGTCAATG 120
Db 61 CCAGAAGTGATGGCAGCTCTCTGGCAGTCCGATCGGATGGAGATGGAGATGCTTGTCAATG 120

Qy 121 AAAGGACCGGTGTTGTTCCATTCCGAGCTACAAAGAAAAAATGTGCATCCAAATCGAG 180
Db 121 AAAGGACCGGTGTTGTTCCATTCCGAGCTACAAAGAAAAAATGTGCATCCAAATCGAG 180

Qy 181 GGGTATATGCCCTTGGATTGATGTTGTCAGCCAGACCTTCCACATTCAGAGACCTT 240
Db 181 GGGTATATGCCCTTGGATTGATGTTGTCAGCCAGACCTTCCACATTCAGAGACCTT 240

Qy 241 AATAAATCATGTTTAAATGCAACACCGGCTACCCCTCTGTGAACAGCAGATTTCTTCGGGTT 300
Db 241 AATAAATCATGTTTAAATGCAACACCGGCTACCCCTCTGTGAACAGCAGATTTCTTCGGGTT 300

Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAGTCAAGTGCAGAACAGAACTCCCAAGGAA 360
Db 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAGTCAAGTGCAGAACAGAACTCCCAAGGAA 360

Qy 361 AAGAAATTCAGAGAAATGAATTTAGCTGTGAGGTATGTGGCGACACATTTAGATCGCT 420
Db 361 AAGAAATTCAGAGAAATGAATTTAGCTGTGAGGTATGTGGCGACACATTTAGATCGCT 420

Qy 421 TTTGATGTTGAGATCCATGAGAACACACAAAGATTTCTTCACTTACGGGTGTAACATG 480
Db 421 TTTGATGTTGAGATCCATGAGAACACACAAAGATTTCTTCACTTACGGGTGTAACATG 480

Qy 481 TCGGAAGAAGMTTSSRRSSAGCCTTGGTTTCTTAAAAATCAATCGCGACACATAATGGC 540
Db 481 TCGGAAGAAGMTTSSRRSSAGCCTTGGTTTCTTAAAAATCAATCGCGACACATAATGGC 540

Qy 541 AAATCGGGGGCCAGAGCAAACTGCAAGGCTTGGAGATGAGTCCAGCAACGATCAAC 600
Db 541 AAATCGGGGGCCAGAGCAAACTGCAAGGCTTGGAGATGAGTCCAGCAACGATCAAC 600

Qy 601 GAGGTCGTCCAGGTGCAACGCGCCGAGAGCATCTCTCTCTTACAAAATCTGATGGTT 660
Db 601 GAGGTCGTCCAGGTGCAACGCGCCGAGAGCATCTCTCTCTTACAAAATCTGATGGTT 660

Qy 661 TGTGCTTCTCTATTTTCCAAATAAGAAAGTCTAAATTGAGCACCGCAAGGTGCAACCAA 720
Db 661 TGTGCTTCTCTATTTTCCAAATAAGAAAGTCTAAATTGAGCACCGCAAGGTGCAACCAA 720
```

```
Db 661 TGTGCTTCTCTATTTTCCAAATAAGAAAGTCTAAATTGAGCACCGCAAGGTGCAACCAA 720
Qy 721 AAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGAGGAATGCCGTCC 780
Db 721 AAAACTGCTTTCGGTACCAGCAGCGCGCAGACAGACTCTCCACAAGAGGAATGCCGTCC 780
Qy 781 TCGAGGAGGAGACTTCTCGCAGTTGTTCAACTTGAGACCAAAATCTCAACCTGAAACGGGG 840
Db 781 TCGAGGAGGAGACTTCTCGCAGTTGTTCAACTTGAGACCAAAATCTCAACCTGAAACGGGG 840
Qy 841 AAGAAAGCTGTGTCAGATGATCCCTCAGCTCGATCCGTTCCACCACTTCCAGGCTTGGCAG 900
Db 841 AAGAAAGCTGTGTCAGATGATCCCTCAGCTCGATCCGTTCCACCACTTCCAGGCTTGGCAG 900
Qy 901 CTGGCTACCAAGGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAAAGGG 960
Db 901 CTGGCTACCAAGGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAAAGGG 960
Qy 961 AGCACCGACAAACGAGATTCCGAGTTCGAGAGAGAGCTTGGAGAAACAAAATAAGGGCAGT 1020
Db 961 AGCACCGACAAACGAGATTCCGAGTTCGAGAGAGAGCTTGGAGAAACAAAATAAGGGCAGT 1020
Qy 1021 TGTGACGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCACGGGGAAGCCCTCC 1080
Db 1021 TGTGACGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCACGGGGAAGCCCTCC 1080
Qy 1081 GTGACGCGGATCCCAAGTTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGC 1140
Db 1081 GTGACGCGGATCCCAAGTTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGC 1140
Qy 1141 GGCAAAAGCTTTTCAAGACCTTACCACAGCTGCTTGTGCACTCCAGGGTCCACAAGAGAAC 1200
Db 1141 GGCAAAAGCTTTTCAAGACCTTACCACAGCTGCTTGTGCACTCCAGGGTCCACAAGAGAAC 1200
Qy 1201 CGGAGGCGCGCGCGGAGTGCACCACCATGCTCTGTGACGGGAGGACCGGGGACGTGT 1260
Db 1201 CGGAGGCGCGCGCGGAGTGCACCACCATGCTCTGTGACGGGAGGACCGGGGACGTGT 1260
Qy 1261 TCTCTGACCTCGCGCGCCCTCTCGATGAAATGGAGCCGTGGATCGAGGGGAAGGTGGT 1320
Db 1261 TCTCTGACCTCGCGCGCCCTCTCGATGAAATGGAGCCGTGGATCGAGGGGAAGGTGGT 1320
Qy 1321 TCTGAAGACGGATCTGAGGATGGGCTTCCGAAAGGAATCCATCTGGATATAAATATGATGAT 1380
Db 1321 TCTGAAGACGGATCTGAGGATGGGCTTCCGAAAGGAATCCATCTGGATATAAATATGATGAT 1380
Qy 1381 GGAGGAAAAATAAACAATCTTACATCTTCAGAGAGTGTAGTTATTTGGGAAAGTTTTTC 1440
Db 1381 GGAGGAAAAATAAACAATCTTACATCTTCAGAGAGTGTAGTTATTTGGGAAAGTTTTTC 1440
Qy 1441 CGTTCAATATTATACCTCAATATTTATCTCAGAACGGATACAGGTGAAAAAACCATACAAA 1500
Db 1441 CGTTCAATATTATACCTCAATATTTATCTCAGAACGGATACAGGTGAAAAAACCATACAAA 1500
Qy 1501 TGTGAATTTTGTGAATATGTCGACGCCAGAGACATCTCTGAGGTATCATCTTGGAGAGA 1560
Db 1501 TGTGAATTTTGTGAATATGTCGACGCCAGAGACATCTCTGAGGTATCATCTTGGAGAGA 1560
Qy 1561 CATCAAGGAAAAAACAACCGATGTTGCTGCTGAAGTCAAGACGATGTTAAAAATCAG 1620
Db 1561 CATCAAGGAAAAAACAACCGATGTTGCTGCTGAAGTCAAGACGATGTTAAAAATCAG 1620
Qy 1621 GACACTGAAGATGCACTATTAAACGCTGACAGTGCAGAAACCAAAAAATTTGAAAAAGATTT 1680
Db 1621 GACACTGAAGATGCACTATTAAACGCTGACAGTGCAGAAACCAAAAAATTTGAAAAAGATTT 1680
Qy 1681 TTTGATGTCGCAAGATGTTTACAGGAGTCCACTCTGCAAGACGCTTAAAGAGATGCTT 1740
Db 1681 TTTGATGTCGCAAGATGTTTACAGGAGTCCACTCTGCAAGACGCTTAAAGAGATGCTT 1740
Qy 1741 TCTGTTTTCAGAAATGTTCTCGGCGAGCGCTCTCTCTCACCAGCACACAAAGATACTCAG 1800
Db 1741 TCTGTTTTCAGAAATGTTCTCGGCGAGCGCTCTCTCTCTCACCAGCACACAAAGATACTCAG 1800
```

```
QY 1801 GATTTCCTAAATAAGTGCAGCTGATGACAGTGTCTGATAAAGTGAATAAATAACCTTACCCCT 1860
DB 1801 GATTTCCTAAATAAGTGCAGCTGATGACAGTGTCTGATAAAGTGAATAAATAACCTTACCCCT 1860
QY 1861 GCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTGTGAAATTCAGGCAAAATAACCTCATC 1920
DB 1861 GCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTGTGAAATTCAGGCAAAATAACCTCATC 1920
QY 1921 TGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAGTACCAACCCATAACCTTGAA 1980
DB 1921 TGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAGTACCAACCCATAACCTTGAA 1980
QY 1981 GTTACGCCCAAGAGAGACAAAGCGAGACCGAGCTGCTGAGATACAGGCAAGTGTG 2040
DB 1981 GTTACGCCCAAGAGAGACAAAGCGAGACCGAGCTGCTGAGATACAGGCAAGTGTG 2040
QY 2041 GATTGTACGAAATACTTTAAATTTATCCGTGGGGCTCTTCAACAATTCGCCGCAATT 2100
DB 2041 GATTGTACGAAATACTTTAAATTTATCCGTGGGGCTCTTCAACAATTCGCCGCAATT 2100
QY 2101 TCTTTGAGTAAAGTTGATTCCTGATCAAGTATCACTGTCCATTTTGTACCTTCAAGCATTT 2160
DB 2101 TCTTTGAGTAAAGTTGATTCCTGATCAAGTATCACTGTCCATTTTGTACCTTCAAGCATTT 2160
QY 2161 TATCCAGAGTTTAAATGATGACACAGAGACTGGAGCATTAATACATCTCGACGTTAT 2220
DB 2161 TATCCAGAGTTTAAATGATGACACAGAGACTGGAGCATTAATACATCTCGACGTTAT 2220
QY 2221 AAAAATGTGCAAAACAAGTCTTGTAGAGTCCAGCTGACCGATGCCCGCACGGTTG 2280
DB 2221 AAAAATGTGCAAAACAAGTCTTGTAGAGTCCAGCTGACCGATGCCCGCACGGTTG 2280
QY 2281 CTGGGAAAGATGTGCTCCCTCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTCTTTC 2340
DB 2281 CTGGGAAAGATGTGCTCCCTCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTCTTTC 2340
QY 2341 CCGGCGCAGTCAAAATCCCTGCTGATCTGCGAAGGGGAGCAGAGCCCTCTCGGCGCAGGC 2400
DB 2341 CCGGCGCAGTCAAAATCCCTGCTGATCTGCGAAGGGGAGCAGAGCCCTCTCGGCGCAGGC 2400
QY 2401 AAGGCCCTCTGACTTTCAGGGATAGACTCTAGCACTTTAGCCCCCAAGTAACTGAAGTCC 2460
DB 2401 AAGGCCCTCTGACTTTCAGGGATAGACTCTAGCACTTTAGCCCCCAAGTAACTGAAGTCC 2460
QY 2461 CACAGACCAACAGCAGATGTGGGGTCCAAAGGGCGCCACACAGGCAACAGCAATCTGAG 2520
DB 2461 CACAGACCAACAGCAGATGTGGGGTCCAAAGGGCGCCACACAGGCAACAGCAATCTGAG 2520
QY 2521 ATGTTTCTTAAACAGTGTGTTTCCCTGACCGGATAGACAAAGAACCCGAGCAAAA 2580
DB 2521 ATGTTTCTTAAACAGTGTGTTTCCCTGACCGGATAGACAAAGAACCCGAGCAAAA 2580
QY 2581 TTGAACCTCTTCCAGTAGTCTCTTCTGAGCCACCTCGGCGAGCAGTAAATCAATGGT 2640
DB 2581 TTGAACCTCTTCCAGTAGTCTCTTCTGAGCCACCTCGGCGAGCAGTAAATCAATGGT 2640
QY 2641 TCCATCGACTACCCGCGCAGAAACAGCAGCCGCTGGGCACTTCCGGGAAGAGACTATTTC 2700
DB 2641 TCCATCGACTACCCGCGCAGAAACAGCAGCCGCTGGGCACTTCCGGGAAGAGACTATTTC 2700
QY 2701 TGTAAATCGAGTGCAGCAATACTGAGCAGAAATTTGGTGAAGCCCTTCCAAAAGAGCTG 2760
DB 2701 TGTAAATCGAGTGCAGCAATACTGAGCAGAAATTTGGTGAAGCCCTTCCAAAAGAGCTG 2760
QY 2761 AGTCCAGGTGTTGCTTGAACCTTGAACCTGACAGCCGCGGCGCAATTAACAGAGGCTAT 2820
DB 2761 AGTCCAGGTGTTGCTTGAACCTTGAACCTGACAGCCGCGGCGCAATTAACAGAGGCTAT 2820
QY 2821 GACCTTCCCAAGTACATATGTCAGAGGCATCATCACTGTTACCGCAGGACTGTGTG 2880
DB 2821 GACCTTCCCAAGTACATATGTCAGAGGCATCATCACTGTTACCGCAGGACTGTGTG 2880
```

```
QY 2881 TATCGTGCAGGCGCTGCTCCCAAAACCAAGGTTCTTGAGCTCCAGCGAGTGCATTTCT 2940
DB 2881 TATCGTGCAGGCGCTGCTCCCAAAACCAAGGTTCTTGAGCTCCAGCGAGTGCATTTCT 2940
QY 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGCTGCTCGGCGCACTTTTACACTTGTGTG 3000
DB 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGCTGCTCGGCGCACTTTTACACTTGTGTG 3000
QY 3001 CCTGCTGTAGTTCAGCATCCAGCTTCGAGCTTTAGAGGTTTGGTGGATGTCACTGCTTA 3060
DB 3001 CCTGCTGTAGTTCAGCATCCAGCTTCGAGCTTTAGAGGTTTGGTGGATGTCACTGCTTA 3060
QY 3061 CTCCCATGAATTAATTTTACTTCTATCCCTTTGAGAGCGCAATCGTGAAGCTACTGAA 3120
DB 3061 CTCCCATGAATTAATTTTACTTCTATCCCTTTGAGAGCGCAATCGTGAAGCTACTGAA 3120
QY 3121 ATAAGCTGTGATTTGATCTGTATCAATAAATATATAGGAATCTGCAAGGAACAATCAGTT 3180
DB 3121 ATAAGCTGTGATTTGATCTGTATCAATAAATATATAGGAATCTGCAAGGAACAATCAGTT 3180
QY 3181 GTGTAA 3186
DB 3181 GTGTAA 3186
```

RESULT 2

```
US-10-087-192-1709
; Sequence 1709, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1709
; LENGTH: 5252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1709
```

```
Query Match 99.9%; Score 3183.6; DB 13; Length 5252;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3180; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGCAATCGAAGTGCAGGAAACATGCCAACTCAATCCCTTTAATGTACATGATGGG 60
DB 25 ATGCAATCGAAGTGCAGGAAACATGCCAACTCAATCCCTTTAATGTACATGATGGG 84
QY 61 CCAAGTGTGAGTGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGATGCGCTTGTCAATG 120
DB 85 CCAAGTGTGAGTGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGATGCGCTTGTCAATG 144
QY 121 AAAGGACCGCTGTTTTCATTCGAGCTACACAAGAAAAAATGTATCAATAAATCGAG 180
DB 145 AAAGGACCGCTGTTTTCATTCGAGCTACACAAGAAAAAATGTATCAATAAATCGAG 204
QY 181 GGGTATATGCTTGGATTTGATGCTTTCGAGCGAGACTTTCACATTCAGAACCTT 240
DB 205 GGGTATATGCTTGGATTTGATGCTTTCGAGCGAGACTTTCACATTCAGAACCTT 264
QY 241 AATAAACATGCTTAAATCAACACCGGCTTACCTCTGTGAACACGAGCTTCTCGGGTT 300
DB 265 AATAAACATGCTTAAATCAACACCGGCTTACCTCTGTGAACACGAGCTTCTCGGGTT 324
```

Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 360
Db 325 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 384
Qy 361 AAGAAATTGCAAGGAAATGAATTTAGTGTGAGGTATGTGGGCGACAGATTTAGAGTGCCT 420
Db 385 AAGAAATTGCAAGGAAATGAATTTAGTGTGAGGTATGTGGGCGACAGATTTAGAGTGCCT 444
Qy 421 TTTGATGTGAGATCCCATCAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 480
Db 445 TTTGATGTGAGATCCCATCAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 504
Qy 481 TCGGGAAGAGTTSRRSSAGCTTGGTTTCTTAAATAATCACATCGCGACACATAATGGC 540
Db 505 TCGGGAAGAGTTCAGAGGCTTGGTTTCTTAAATAATCACATCGCGACACATAATGGC 564
Qy 541 AAATCGGGGCGAGAACTGACAGCAAGCAAGCTTTGGAGAGTAGTCCAGCAACGATCAAC 600
Db 565 AAATCGGGGCGAGAACTGACAGCAAGCAAGCTTTGGAGAGTAGTCCAGCAACGATCAAC 624
Qy 601 GAGTCTGTCAGGTGCACGCGCGAGAGCATCTCCTCTCTTACAAATCTGCATGGTT 660
Db 625 GAGTCTGTCAGGTGCACGCGCGAGAGCATCTCCTCTCTTACAAATCTGCATGGTT 684
Qy 661 TGTGCTTCTTAATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAA 720
Db 685 TGTGCTTCTTAATTTCCAAATANAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAA 744
Qy 721 AAAAATGCTTTTGGTACACAGAGCGCGACAGACTCTCCACAAAGGAGGAATGCCGTCC 780
Db 745 AAAAATGCTTTTGGTACACAGAGCGCGACAGACTCTCCACAAAGGAGGAATGCCGTCC 804
Qy 781 TCGAGGAGGACTTCTCCTCAGTGTGTTCACTTGAGACCAAAATCTCACCCCTGMAACCGGG 840
Db 805 TCGAGGAGGACTTCTCCTCAGTGTGTTCACTTGAGACCAAAATCTCACCCCTGMAACCGGG 864
Qy 841 AAGAACTGTCAGATGATCCCTCAGTCCGATCCGTTTCCACACCTTCCAGCTTGGCAG 900
Db 865 AAGAACTGTCAGATGATCCCTCAGTCCGATCCGTTTCCACACCTTCCAGCTTGGCAG 924
Qy 901 CTGGCTACCAAGGAAAGTTCGCAATTTGCCAAGAAAGTGAAGGAATCGGGGCAAGAGGG 960
Db 925 CTGGCTACCAAGGAAAGTTCGCAATTTGCCAAGAAAGTGAAGGAATCGGGGCAAGAGGG 984
Qy 961 AGCACGACACGAGATTCGAGTTCGAGAGGAGCTTGAGAAACAAATAAGGCGAGT 1020
Db 985 AGCACGACACGAGATTCGAGTTCGAGAGGAGCTTGAGAAACAAATAAGGCGAGT 1044
Qy 1021 TGTGAGGCTCTCGCAAGAGAAAGAGAGTGCAAACACTCCACGCGGAGCGCCCTCC 1080
Db 1045 TGTGAGGCTCTCGCAAGAGAAAGAGAGTGCAAACACTCCACGCGGAGCGCCCTCC 1104
Qy 1081 GTGGAACGGGATCCCAAGTTACCAAGTAGCAAGAGAGAGCCACCTCACTGCTCCAGTGC 1140
Db 1105 GTGGAACGGGATCCCAAGTTACCAAGTAGCAAGAGAGAGCCACCTCACTGCTCCAGTGC 1164
Qy 1141 GGCAGAGCTTTCAGAACTTACACAGCTGCTTCCACTCCAGGTCACAGAGGAGC 1200
Db 1165 GGCAGAGCTTTCAGAACTTACACAGCTGCTTCCACTCCAGGTCACAGAGGAGC 1224
Qy 1201 CGGAGGCGGCGGAGTTCGCGCCACCATGTCTGGAACGGGAGCAGCCGGGACGTGT 1260
Db 1225 CGGAGGCGGCGGAGTTCGCGCCACCATGTCTGGAACGGGAGCAGCCGGGACGTGT 1284
Qy 1261 TCTCTGACCTCGCCGCCCTCTGGATGAAATAGAGCGGTGGATCGAGGGGAAGGTGT 1320
Db 1285 TCTCTGACCTCGCCGCCCTCTGGATGAAATAGAGCGGTGGATCGAGGGGAAGGTGT 1344
Qy 1321 TCTGAGAGCGGATCTGAGATGGGCTTCCGAGGAGTCCATCTCGATATAAATATGAT 1380
Db 1345 TCTGAGAGCGGATCTGAGATGGGCTTCCGAGGAGTCCATCTCGATATAAATATGAT 1404
Qy 1381 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTC 1440

Db 1405 GGAGGAAAAATAAACATCTTACATCTTCAAGAGGTGTAGTTATTGTGGAAAGTTTTTC 1464
Qy 1441 CGTTCAAAATTTATACCTCAATATTCATCTCAGAACGCATACAGGTGAAACCATACAAA 1500
Db 1465 CGTTCAAAATTTATACCTCAATATTCATCTCAGAACGCATACAGGTGAAACCATACAAA 1524
Qy 1501 TGTGAAATTTTGTGAATATGCTGCAGCCGAGAACATCTCTGAGGTATCACTTGGAGAGA 1560
Db 1525 TGTGAAATTTTGTGAATATGCTGCAGCCGAGAACATCTCTGAGGTATCACTTGGAGAGA 1584
Qy 1561 CATCAAGGAAAAACAAACCGATGTTGCTGCTGAAAGTCAAGAACGATGTTAAATAATCAG 1620
Db 1585 CATCAAGGAAAAACAAACCGATGTTGCTGCTGAAAGTCAAGAACGATGTTAAATAATCAG 1644
Qy 1621 GACACTGAAGTCACTATTAAACCGCTGACAGTGCAGAACCAAAATTTTGAAGAATTT 1680
Db 1645 GACACTGAAGTCACTATTAAACCGCTGACAGTGCAGAACCAAAATTTTGAAGAATTT 1704
Qy 1681 TTTGATGTCGCAAGATGTTACAGGAGTCCACTGCAAGCAGCTTAAAGAGATGCCT 1740
Db 1705 TTTGATGTCGCAAGATGTTACAGGAGTCCACTGCAAGCAGCTTAAAGAGATGCCT 1764
Qy 1741 TCTGTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACAGCACACAAAGATCTCAG 1800
Db 1765 TCTGTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACAGCACACAAAGATCTCAG 1824
Qy 1801 GATTTCCATAAATGTCAGCTGATGACAGTCTCATAAAGTGAATATAAAACCTTACCCCT 1860
Db 1825 GATTTCCATAAATGTCAGCTGATGACAGTCTCATAAAGTGAATATAAAACCTTACCCCT 1884
Qy 1861 GCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTGTGAAAATCTCAGGCAAAATTAACCTCATC 1920
Db 1885 GCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTGTGAAAATCTCAGGCAAAATTAACCTCATC 1944
Qy 1921 TGTAGAACCAAGGCGGATGTTTAAATTTATFCCGTGGGGCTTTTCAAAATTTGCCGCAAT 2004
Db 1945 TGTAGAACCAAGGCGGATGTTTAACTCTCTCCGATGGCAGTAGCACCAATTAACCTTGA 2040
Qy 1981 GTTAGCCCCCAAGAGAGCAACCGAGACCGCAGCTGACTGCAGATACAGGCCAGTGTG 2040
Db 2005 GTTAGCCCCCAAGAGAGCAACCGAGACCGCAGCTGACTGCAGATACAGGCCAGTGTG 2064
Qy 2041 GATTTGTCACGAAAAACCTTTTAAATTTATFCCGTGGGGCTTTTCAAAATTTGCCGCAAT 2100
Db 2065 GATTTGTCACGAAAAACCTTTTAAATTTATFCCGTGGGGCTTTTCAAAATTTGCCGCAAT 2124
Qy 2101 TCTTTGAGTAAAAAGTTTGAATTTCAAGTATCACTGTCTCAATTTTGTACCTTCAAGACATTT 2160
Db 2125 TCTTTGAGTAAAAAGTTTGAATTTCAAGTATCACTGTCTCAATTTTGTACCTTCAAGACATTT 2184
Qy 2161 TATCCAGAGTTTTTAAATGATGCACAGAGCTGGAGCATATAATCAATCTTGACGTTTCAAT 2220
Db 2185 TATCCAGAGTTTTTAAATGATGCACAGAGCTGGAGCATATAATCAATCTTGACGTTTCAAT 2244
Qy 2221 AAAAACTGTCGAAAAACAGTCTTGTAGAAAGTGCAGTACCGGATGCCGCGCAGCGTTG 2280
Db 2245 AAAAACTGTCGAAAAACAGTCTTGTAGAAAGTGCAGTACCGGATGCCGCGCAGCGTTG 2304
Qy 2281 CTGGGAAAAAGATGTGCTCTCTCTCTAGTTTTTGTATAAACCCAGCCCAAGTCTGCTTTC 2340
Db 2305 CTGGGAAAAAGATGTGCTCTCTCTCTAGTTTTTGTATAAACCCCAAGTCTGCTTTC 2364
Qy 2341 CGGGCGAGTCCAAATCCCTGCCATCTCGAGGAGGAGAGAGCCCTCTTGGGGCAGGC 2400
Db 2365 CGGGCGAGTCCAAATCCCTGCCATCTCGAGGAGGAGAGAGCCCTCTTGGGGCAGGC 2424
Qy 2401 AAGGCCCTCTGACTTTCAGGGATAGACTCTAGCACTTTTAGCCCCCAAGTAACCTGAAGTCC 2460
Db 2425 AAGGCCCTCTGACTTTCAGGGATAGACTCTAGCACTTTTAGCCCCCAAGTAACCTGAAGTCC 2484
Qy 2461 CACAGACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACAGGCCAACAGCAATCTGAG 2520

Db 2485 CACAGACCACAGAGAAATGTGGGGTCCAAAGGGCGCCACACAGGCAACAGCAATCTGAG 2544
Qy 2521 ATGTTTCTTAAACCAAGTGTTCCTCCCTGACCGGTAAGACAAAAGACCCGAGACAAA 2580
Db 2545 ATGTTTCTTAAACCAAGTGTTCCTCCCTGACCGGTAAGACAAAAGACCCGAGACAAA 2604
Qy 2581 TTGAACCTCTTCCAGTAGTCTCTCTCAGCCACCTCGGAGAGTAACATCAATGCT 2640
Db 2605 TTGAACCTCTTCCAGTAGTCTCTCTCAGCCACCTCGGAGAGTAACATCAATGCT 2664
Qy 2641 TCCATCGACTACCCCGCAAGACAGACCGCGTGGGCACTTCCGGGAGAGACTATTTC 2700
Db 2665 TCCATCGACTACCCCGCAAGACAGACCGCGTGGGCACTTCCGGGAGAGACTATTTC 2724
Qy 2701 TGTATCGAGTGCAGCAGATCTGACGAGAGATTTGGTGAGCCCTTCCAAAAGACTG 2760
Db 2725 TGTATCGAGTGCAGCAGATCTGACGAGAGATTTGGTGAGCCCTTCCAAAAGACTG 2784
Qy 2761 AAGTCCAGGCTGTTGGCCCTTACAGTGTACACAGCCCGGGCCAAATTACAGAGAGGCTAT 2820
Db 2785 AAGTCCAGGCTGTTGGCCCTTACAGTGTACACAGCCCGGGCCAAATTACAGAGAGGCTAT 2844
Qy 2821 GACCTTCCCAAGTACCATATGCTCAGAGGCATCACATCACTGTTACCGAGGACTGTGTG 2880
Db 2845 GACCTTCCCAAGTACCATATGCTCAGAGGCATCACATCACTGTTACCGAGGACTGTGTG 2904
Qy 2881 TATCGTGCAGCGCTGCTCCCAACCAAGGTTCTGAGTCTCAGGAGGTCGATTC 2940
Db 2905 TATCGTGCAGCGCTGCTCCCAACCAAGGTTCTGAGTCTCAGGAGGTCGATTC 2964
Qy 2941 CCAATGTCGACTGTTGAGAGCCCTATGAGTGTGCTCGGCGCACTTTACATTTGCTG 3000
Db 2965 CCAATGTCGACTGTTGAGAGCCCTATGAGTGTGCTCGGCGCACTTTACATTTGCTG 3024
Qy 3001 CTGCTGTGAGTGCAGAGTCCAGCTCGAGCTTAGAAGTCTTGGTGGATGTGAGTCTTA 3060
Db 3025 CTGCTGTGAGTGCAGAGTCCAGCTCGAGCTTAGAAGTCTTGGTGGATGTGAGTCTTA 3084
Qy 3061 CTCCCAGTAATTAATTTACTTCTTCTGAGAGCGGAATGGTGAAGCTACTGAA 3120
Db 3085 CTCCCAGTAATTAATTTACTTCTTCTGAGAGCGGAATGGTGAAGCTACTGAA 3144
Qy 3121 ATAAGCTGTGATGTACTGTACATAAATATAGGAATCTGCAAGGAACTACATGCTT 3180
Db 3145 ATAAGCTGTGATGTACTGTACATAAATATAGGAATCTGCAAGGAACTACATGCTT 3204
Qy 3181 GTGTAA 3186
Db 3205 GTGTAA 3210

RESULT 3

US-10-177-293-505
; Sequence 505, Application US/10177293
; Publication No. US200301241281
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul

; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 505
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-505

Query Match 95.4%; Score 3040.6; DB 15; Length 5632;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;

Qy 1 ATGCATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAAATGTACATGATGGG 60
Db 272 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAAATGTACATGATGGG 331
Qy 61 CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGSCCTTGTCAATG 120
Db 332 CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGSCCTTGTCAATG 391
Qy 121 AAAGGACCGCTGTTGTTTCCATTCGAGCTACACAAGAAAAAATGTTCATCCAAATCGAG 180
Db 392 AAAGGACCGCTGTTGTTTCCATTCGAGCTACACAAGAAAAAATGTTCATCCAAATCGAG 451
Qy 181 GGGTATATGCCCTGGATTGATGTTCTGCGAGCAGACCTTCACACATTCAGAACCTT 240
Db 452 GGGTATATGCCCTGGATTGATGTTCTGCGAGCAGACCTTCACACATTCAGAACCTT 511
Qy 241 AATAAACATGCTTAAATGCAACACCGGCTACCCCTCTGTAACACAGCAGTCTTCGGGTT 300
Db 512 AATAAACATGCTTAAATGCAACACCGGCTACCCCTCTGTAACACAGCAGTCTTCGGGTT 571
Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCAGAACAGAACCTTCCCAAGAA 360
Db 572 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCAGAACAGAACCTTCCCAAGAA 631
Qy 361 AAGAAATTCAGGAAATGAATTTAGCTGTGAGGTATGTGGGAGACATTTAGAGTCGT 420
Db 632 AAGAAATTCAGGAAATGAATTTAGCTGTGAGGTATGTGGGAGACATTTAGAGTCGT 691
Qy 421 TTTGATGTTGAGATCCACATGAGAACACACAAAGATCTTTTTCATTCAGGGTGAACATG 480
Db 692 TTTGATGTTGAGATCCACATGAGAACACACAAAGATCTTTTTCATTCAGGGTGAACATG 751
Qy 481 TCGGAAAGAAAGTTSSRSSAGCCTTGGTTTCTTAAAAATCATATCGGAGACATATGCG 540
Db 752 TCGGAAAGAAAGTTCAAGGAGCCTTGGTTTCTTAAAAATCATATCGGAGACATATGCG 811
Qy 541 AAATCGGGGGCCAGAGCAAACTGAGCAAGAGCTTTGGAGATGTATCCAGCAACGATCAAC 600
Db 812 AAATCGGGGGCCAGAGCAAACTGAGCAAGAGCTTTGGAGATGTATCCAGCAACGATCAAC 871
Qy 601 GAGTCTGTGAGTGCAGCGCGGAGAGCATCTCTCTCTTCAAAAATCTGCATGCTT 660
Db 872 GAGTCTGTGAGTGCAGCGCGGAGAGCATCTCTCTCTTCAAAAATCTGCATGCTT 931

| | | | |
|----|------|---|------|
| Qy | 661 | TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGACACCAA | 720 |
| Db | 932 | TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGACACCAA | 991 |
| Qy | 721 | AAAACTGCTTTTCGGTACAGCAGCGCGCAGACAGACTCTCCACAGGAGGAATGCCGTCC | 780 |
| Db | 992 | AAAACTGCTTTTCGGTACAGCAGCGCGCAGACAGACTCTCCACAGGAGGAATGCCGTCC | 1051 |
| Qy | 781 | TCGAGGAGGAGCTTCCTGCAAGTTGTTTCAACTTCGAGACCAAAATCTCACCTGAAACGGGG | 840 |
| Db | 1052 | TCGAGGAGGAGCTTCCTGCAAGTTGTTTCAACTTCGAGACCAAAATCTCACCTGAAACGGGG | 1111 |
| Qy | 841 | AAGAAGCCTGTACAGTGCATCCCTCAGCTCGATCCGTTTCCACCTTCCAGGCTTGGCAG | 900 |
| Db | 1112 | AAGAAGCCTGTACAGTGCATCCCTCAGCTCGATCCGTTTCCACCTTCCAGGCTTGGCAG | 1171 |
| Qy | 901 | CTGCTCTACAAAGGAAAAGTTGCCAATTGGCAAGAAGTGAAGGAATCGGGCACAAGGG | 960 |
| Db | 1172 | CTGCTCTACAAAGGAAAAGTTGCCAATTGGCAAGAAGTGAAGGAATCGGGCACAAGGG | 1231 |
| Qy | 961 | AGCACGACAACGACGATTTCGAGTTCGAGACGAGGAGCTTCGAGAAAACAAATATAGGCGAGT | 1020 |
| Db | 1232 | AGCACGACAACGACGATTTCGAGTTCGAGACGAGGAGCTTCGAGAAAACAAATATAGGCGAGT | 1291 |
| Qy | 1021 | TGTGCGGCTCTCGCAAGAGAAAGAGAGTGCACACACTCCACGCGCAAGCGCCCTCC | 1080 |
| Db | 1292 | TGTGCGGCTCTCGCAAGAGAAAGAGAGTGCACACACTCCACGCGCAAGCGCCCTCC | 1351 |
| Qy | 1081 | GTGACGCGGATCCCAAGTTTACCCAGTAGCAAGAGAGAGCCCACTCACTGCTCCGAGTGC | 1140 |
| Db | 1352 | GTGACGCGGATCCCAAGTTTACCCAGTAGCAAGAGAGAGCCCACTCACTGCTCCGAGTGC | 1411 |
| Qy | 1141 | GGCAAGCTTTTCAGAACCTTACCAACAGTGGTCTTTGCACTTCAGGTTCCACAGAAGGAC | 1200 |
| Db | 1412 | GGCAAGCTTTTCAGAACCTTACCAACAGTGGTCTTTGCACTTCAGGTTCCACAGAAGGAC | 1471 |
| Qy | 1201 | CGGAGGCGCGCGGAGTGCSCCACCATGTCTGTGACGGGAGGAGCGCGGAGCTGT | 1260 |
| Db | 1472 | CGGAGGCGCGCGGAGTGCSCCACCATGTCTGTGACGGGAGGAGCGCGGAGCTGT | 1531 |
| Qy | 1261 | TCTCCTGACCTTCGCGCGCCCTCTTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGTGGT | 1320 |
| Db | 1532 | TCTCCTGACCTTCGCGCGCCCTCTTGGATGAAAATGGAGCCGTGGATCGAGGGGAAGTGGT | 1591 |
| Qy | 1321 | TCTGAAGACGGATCTGAGGATGGGCTTCCGAGGAATCCATCTGGATATAAATATGATGAT | 1380 |
| Db | 1592 | TCTGAAGACGGATCTGAGGATGGGCTTCCGAGGAATCCATCTGGATATAAATATGATGAT | 1651 |
| Qy | 1381 | GGAGGAAAATTAACAATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTC | 1440 |
| Db | 1652 | GGAGGAAAATTAACAATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTC | 1711 |
| Qy | 1441 | CGTTCAAATTAATACCTCAATATTTCATCTCAGAACGCATACAGGTGAAAACCAACAAA | 1500 |
| Db | 1712 | CGTTCAAATTAATACCTCAATATTTCATCTCAGAACGCATACAGGTGAAAACCAACAAA | 1771 |
| Qy | 1501 | TGTGAATTTTGTGAATATGCTGCAGCCCCAGAGAATCTCTGTGAGGTATCATCTTGGAGAGA | 1560 |
| Db | 1772 | TGTGAATTTTGTGAATATGCTGCAGCCCCAGAGAATCTCTGTGAGGTATCATCTTGGAGAGA | 1831 |
| Qy | 1561 | CATCAAGAGGAAAACAAACCGATGTTCCTGCTGAGTCAAGACGATGTTAAAAATCAG | 1620 |
| Db | 1832 | CATCAAGAGGAAAACAAACCGATGTTCCTGCTGAGTCAAGACGATGTTAAAAATCAG | 1891 |
| Qy | 1621 | GACACTGAAGATGCACATTAATAACCGCTGCAGCTGCGCAACCAAAAAATTTGAAAAGATTT | 1680 |
| Db | 1892 | GACACTGAAGATGCACATTAATAACCGCTGCAGCTGCGCAACCAAAAAATTTGAAAAGATTT | 1951 |
| Qy | 1681 | TTTGATGTCGCAAGATGTTTACAGGCAAGTCCACTCGTCAAAAGCAGCTTAAAGGAGATCCCT | 1740 |
| Db | 1952 | TTTGATGTCGCAAGATGTTTACAGGCAAGTCCACTCGTCAAAAGCAGCTTAAAGGAGATCCCT | 2011 |

| | | | |
|----|------|--|------|
| Qy | 1741 | TCGTGTTTTTCAGAAATGTTCTCTGGGCAGCGGTCTCTCTCAACAGCACACAAGATACTCTAG | 1800 |
| Db | 2012 | TCGTGTTTTTCAGAAATGTTCTGGGCAGCGGTCTCTCTCACACAGCACACAAGATACTCTAG | 2071 |
| Qy | 1801 | GATTTCCATAAAAAATCCAGCTGATGACGTGCTGTATATAAGTGAATAAAAAACCTCAACCCCT | 1860 |
| Db | 2072 | GATTTCCATAAAAAATGGAGCTGATGACGTGCTGTATATAAGTGAATAAAAAACCTCAACCCCT | 2131 |
| Qy | 1861 | GCTTTACTCTGACCTGTTAAAAAGAGATCAGCAGTTTGAACCTCAGGCCAAATAACCTCATC | 1920 |
| Db | 2132 | GCTTTACTGACCTGTTAAAAAGAGATCAGCAGTTTGAACCTCAGGCCAAATAACCTCATC | 2191 |
| Qy | 1921 | TGTAGAACCAAGCGCGATGTTACTCTCTCTCGGATGGCAGTACCAACCCATAACCTTGAA | 1980 |
| Db | 2192 | TGTAGAACCAAGCGCGATGTTACTCTCTCTCGGATGGCAGTACCAACCCATAACCTTGAA | 2251 |
| Qy | 1981 | GTTAGCCCCAAGAGAACAAACGGAGACGGCAGCTGACTGACAGATACAGGCCCAAGTGTG | 2040 |
| Db | 2252 | GTTAGCCCCAAGAGAACAAACGGAGACGGCAGCTGACTGACAGATACAGGCCCAAGTGTG | 2311 |
| Qy | 2041 | GATTTGCACCAAAAAACCTTTTAAATTTATCCGTGGGGCTCTTCACAAATTTGCCCGGCAATT | 2100 |
| Db | 2312 | GATTTGCACCAAAAAACCTTTTAAATTTATCCGTGGGGCTCTTCACAAATTTGCCCGGCAATT | 2371 |
| Qy | 2101 | TCCTTGAGTAAAAAGTTTGATTTCCAAGTATCACCTGTCCAATTTTGTACTTTCAAGACATTT | 2160 |
| Db | 2372 | TCCTTGAGTAAAAAGTTTGATTTCCAAGTATCACCTGTCCAATTTTGTACTTTCAAGACATTT | 2431 |
| Qy | 2161 | TATCCAGAAAGTTTAAATGATGACACAGAGACTGGAGCATTAATAACATCTCTGACGTTTCAT | 2220 |
| Db | 2432 | TATCCAGAAAGTTTAAATGATGACACAGAGACTGGAGCATTAATAACATCTCTGACGTTTCAT | 2491 |
| Qy | 2221 | AAAACTGTCGAAACAAAGTCTCTTGTAGAAAGTCGAGCTACCGGATGCCCGCCAGCGTTG | 2280 |
| Db | 2492 | AAAACTGTCGAAACAAAGTCTCTTGTAGAAAGTCGAGCTACCGGATGCCCGCCAGCGTTG | 2551 |
| Qy | 2281 | CTGGGAAAAGATGTGCCCTCCCTCTCTAGTPTTCTGTAAACCCAAAGCCCAAGTCTGCTTTC | 2340 |
| Db | 2552 | CTGGGAAAAGATGTGCCCTCCCTCTCTAGTPTTCTGTAAACCCAAAGCCCAAGTCTGCTTTC | 2611 |
| Qy | 2341 | CCGGCGAGTCCAAAATCCCTGCCATCTGCGAAGGGAGCAGAGCCCTCCTGGGCCAGGC | 2400 |
| Db | 2612 | CCGGCGAGTCCAAAATCCCTGCCATCTGCGAAGGGAGCAGAGCCCTCCTGGGCCAGGC | 2461 |
| Qy | 2401 | AAGGCCCTCTGACTTTCAGGGATAGACTCTAGCAGCTTTAGCCCAAGTAACTTGAAGTCC | 2460 |
| Db | 2672 | AAGGCCCTCTGACTTTCAGGGATAGACTCTAGCAGCTTTAGCCCAAGTAACTTGAAGTCC | 2731 |
| Qy | 2461 | CACAGACACAGCAGAAATGTGGGGGTCCAAAGGGCCGCCACAGGCCAACAGCAATCTGAG | 2520 |
| Db | 2732 | CACAGACACAGCAGAAATGTGGGGGTCCAAAGGGCCGCCACAGGCCAACAGCAATCTGAG | 2791 |
| Qy | 2521 | ATGTTTCTTAAAAACCAAGTGTTCCTCTGCAACCGGATAAGACAAAAAGACCCGAGACAAA | 2580 |
| Db | 2792 | ATGTTTCTTAAAAACCAAGTGTTCCTCTGCAACCGGATAAGACAAAAAGACCCGAGACAAA | 2851 |
| Qy | 2581 | TTGAAACCTCTTCCAGTGTCTCTCTCAGCCCAACCTCGGCAGCAGTACATCAATGTT | 2640 |
| Db | 2852 | TTGAAACCTCTTCCAGTGTCTCTCTCAGCCCAACCTCGGCAGCAGTACATCAATGTT | 2911 |
| Qy | 2641 | TCCATCTGACTACCCCGCAAGAACGACAGCCCGTGGGCACCTCCGGGAAAGAGACTATTTTC | 2700 |
| Db | 2912 | TCCATCTGACTACCCCGCAAGAACGACAGCCCGTGGGCACCTCCGGGAAAGAGACTATTTTC | 2971 |
| Qy | 2701 | TGTAATCGGAGTGCAGCAATACTCTGACAGCAGAAATTTGGTGAGGCCCTTCCAAAAAGACTG | 2760 |
| Db | 2972 | TGTAATCGGAGTGCAGCAATACTCTGACAGCAGAAATTTGGTGAGGCCCTTCCAAAAAGACTG | 3031 |
| Qy | 2761 | AAGTCCAGCGTGGTTGCCCTTGAGCTTGACAGCCCGGGGCCAATTAACAGAGAGGCTAT | 2820 |
| Db | 3032 | AAGTCCAGCGTGGTTGCCCTTGAGCTTGACAGCCCGGGGCCAATTAACAGAGAGGCTAT | 3091 |
| Qy | 2821 | GACCTTCCCAAGTACCATATGTTTCAGAGGCATCACATCACTGTGTACCGCAGGACTGTGTG | 2880 |

Db 3092 GACCTTCCCAAGTACCATATGTCAGAGGCATCATCACTGTACCGCAGGACTGTGTG 3151
Qy 2881 TATCGTGCAGCGCTGCTCCAAAACCAAGGTTCTGAGCTCCAGCGAGGTGCAATCT 2940
Db 3152 TATCGTGCAGCGCTGCTCCAAAACCAAGGTTCTGAGCTCCAGCGAGGTGCAATCT 3211
Qy 2941 CCAAAATGTGCTGACTGTTTCAGAAAGCCCATGTGTGCTCCGGGCCACTTTACACTTGTGTG 3000
Db 3212 CCAAAATGTGCTGACTGTTTCAGAAAGCCCATGTGTGCTCCGGGCCACTTTACACTTGTGTG 3271
Qy 3001 CTGCTGTGTGTCAGCTGTTTCAGAAAGCCCATGTGTGCTCCGGGCCACTTTACACTTGTGTG 3034
Db 3272 CTGCTGTGTGTCAGCTGTTTCAGAAAGCCCATGTGTGCTCCGGGCCACTTTACACTTGTGTG 3331
Qy 3035 ----- 3034
Db 3332 CACTTATCTAACAGCATGSCACAAAAGAGAAACTATGAGAAATTTATTGGGAATGCACAT 3391
Qy 3035 -----AAGGTCTTGGTGG 3047
Db 3392 TATGACCAAAATGACAAAAAACTTGATTCTACTAATTTAGGGGGAAAAAAGGTCITGGTGG 3451
Qy 3048 ATGTCAGTGTCTTACTCCCATGAAATTTAAATTTTACTTCTTCTGAGAAAGCAATGCT 3107
Db 3452 ATGTCAGTGTCTTACTCCCATGAAATTTAAATTTTACTTCTTCTGAGAAAGCAATGCT 3511
Qy 3108 GAAAGCTACTGAAATTAAGCTGTGATGTGCTGTACATATAAACAATATGAGGAATCTGCAAG 3167
Db 3512 GAAAGCTACTGAAATTAAGCTGTGATGTGCTGTACATATAAACAATATGAGGAATCTGCAAG 3571
Qy 3168 GAACACTACAGTTGTGTAA 3186
Db 3572 GAACACTACAGTTGTGTAA 3590

RESULT 4

US-10-087-192-1708
; Sequence 1708, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIORITY FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1708
; LENGTH: 36022
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1708

Query Match 49.0%; Score 1561.2; DB 13; Length 36022;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGAGACATCTCT 1541
Db 15816 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGAGACATCTCT 15875
Qy 1542 GAGGTATCACTTGGAGAGACATCAACAGGAAAAACCAACCGATGTTGCTGTGAAGTCAA 1601
Db 15876 GAGGTATCACTTGGAGAGACATCAACAGGAAAAACCAACCGATGTTGCTGTGAAGTCAA 15935
Qy 1602 GAACGATGGTAAAAATCAGGACATGAAGATGCACATTTAACCCGCTGACAGTGGCCAAAC 1661

Db 15936 GAACGATGGTAAAAATCAGGACATGAAGATGCACATTTAAACCGCTGACAGTGGCCAAAC 15995
Qy 1662 CAAAAATTTGAAAGATTTTTTGTGATGGTGCACCAAGATGTTTACAGCAGTCCACCTGCAAA 1721
Db 15996 CAAAAATTTGAAAGATTTTTTGTGATGGTGCACCAAGATGTTTACAGCAGTCCACCTGCAAA 16055
Qy 1722 CGAGCTTAAGGAGATGCTTCTGTTTTCAGAAATGTTTTCGGGCGAGGCTGCTCTCACC 1781
Db 16056 CGAGCTTAAGGAGATGCTTCTGTTTTCAGAAATGTTTTCGGGCGAGGCTGCTCTCACC 16115
Qy 1782 AGCACACAAAGATACTCAGGATTTTCCATAAAATGACAGCTGATGACAGTGTGATAAAGT 1841
Db 16116 AGCACACAAAGATACTCAGGATTTTCCATAAAATGACAGCTGATGACAGTGTGATAAAGT 16175
Qy 1842 GAATAAAAACCTTACCTCTGCTTACTCTGGAACCTGTTTAAAAAAGAGATCAGCAGTGTGAAAC 1901
Db 16176 GAATAAAAACCTTACCTCTGCTTACTCTGGAACCTGTTTAAAAAAGAGATCAGCAGTGTGAAAC 16235
Qy 1902 TCAGGCAATTAACCTCTGCTGATGAAACCAAGCGGATGTTTACTCTCTCGGATGGGAG 1961
Db 16236 TCAGGCAATTAACCTCTGCTGATGAAACCAAGCGGATGTTTACTCTCTCGGATGGGAG 16295
Qy 1962 TACCACCCATTAACCTTGAAGTTTAGCCCCCAAGAGAGAAAGCAACCGAGACCGCAGCTGACTG 2021
Db 16296 TACCACCCATTAACCTTGAAGTTTAGCCCCCAAGAGAGAAAGCAACCGAGACCGCAGCTGACTG 16355
Qy 2022 CAGATACAGGCCAAGTGTGGAATGTCAGAAAAAATCTTTTAAATTTTATCCGTGGGGGCTCT 2081
Db 16356 CAGATACAGGCCAAGTGTGGAATGTCAGAAAAAATCTTTTAAATTTTATCCGTGGGGGCTCT 16415
Qy 2082 TCACAAATGGCCCGCAATTTCTTTGAGTAAAAAGTTTGAATCCAGATATCACTGTCCATT 2141
Db 16416 TCACAAATGGCCCGCAATTTCTTTGAGTAAAAAGTTTGAATCCAGATATCACTGTCCATT 16475
Qy 2142 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACACAGAGACTGAGACATAA 2201
Db 16476 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACACAGAGACTGAGACATAA 16535
Qy 2202 ATACAATCTGACGTTCTATAAAAAATCTGCGAAAAAGTCTTGTCTTAGAAGTTCGACGTAC 2261
Db 16536 ATACAATCTGACGTTCTATAAAAAATCTGCGAAAAAGTCTTGTCTTAGAAGTTCGACGTAC 16595
Qy 2262 CGATGCCCCCGCAGGTTGCTGGGAAAAAGATGTGCTCCCTCTCTAGTCTTCTGTAACCC 2321
Db 16596 CGATGCCCCCGCAGGTTGCTGGGAAAAAGATGTGCTCCCTCTCTAGTCTTCTGTAACCC 16655
Qy 2322 CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCGCATCTGCGAAGGGGAAGCA 2381
Db 16656 CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCGCATCTGCGAAGGGGAAGCA 16715
Qy 2382 GAGCCCTCTCTGGGCCAGGCAAGGCCCTCTGACTTTAGGAGATAGACTCTAGCATTAGC 2441
Db 16716 GAGCCCTCTCTGGGCCAGGCAAGGCCCTCTGACTTTAGGAGATAGACTCTAGCATTAGC 16775
Qy 2442 CCAAGTAACTGAAATCCCAAGACACAGAGATGTTGGGGTCCAAAGGGGGCCGCCAC 2501
Db 16776 CCAAGTAACTGAAATCCCAAGACACAGAGATGTTGGGGTCCAAAGGGGGCCGCCAC 16835
Qy 2502 CAGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTCTTCCCTGCGACCGGATAAGAC 2561
Db 16836 CAGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTCTTCCCTGCGACCGGATAAGAC 16895
Qy 2562 AAAAAGACCCGAGACAAAAATTTGAAACCTTCTTCCAGTAGCTCTCTCAGCCCCACCCTCGG 2621
Db 16896 AAAAAGACCCGAGACAAAAATTTGAAACCTTCTTCCAGTAGCTCTCTCAGCCCCACCCTCGG 16955
Qy 2622 CAGCAGTAAATCAATGTTCCATCGACTACCCCGCAAGAAACGACACCCCGTGGGCACC 2681
Db 16956 CAGCAGTAAATCAATGTTCCATCGACTACCCCGCAAGAAACGACACCCCGTGGGCACC 17015
Qy 2682 TCCGGGAGAGACTATTTCTGTAATCCGAGTGCAGCAATACTCGCAGAGAAATTTGGTGA 2741

Db 17016 TCCGGGAAGAGACTATTTCCTGTAATCGGAGTGCACGCAATACTCGACAGAAATTTGGTGA 17075
Qy 2742 GCCCTTCCAAAAGACTGAAGTCAGCGTGGTGGCCCTTGACGTTGACCAAGCCGGGGC 2801
Db 17076 GCCCTTCCAAAAGACTGAAGTCAGCGTGGTGGCCCTTGACGTTGACCAAGCCGGGGC 17135
Qy 2802 CAATTACAGAAGAGCTATGACCTTCCCAAGTACCATATGCTCAGAGCATCACATCACT 2861
Db 17136 CAATTACAGAAGAGCTATGACCTTCCCAAGTACCATATGCTCAGAGCATCACATCACT 17195
Qy 2862 GTTACCGCAGACTGTGTGATATCGTGCAGCGCGTGCCTCCAAAACCAAGTTCCTGAG 2921
Db 17196 GTTACCGCAGACTGTGTGATATCGTGCAGCGCGTGCCTCCAAAACCAAGTTCCTGAG 17255
Qy 2922 CTCAGCGAGGTGCATCTCCAAATGTGCTCACTGTTTCAGAAAGCCCTATGCTGCTCCGG 2981
Db 17256 CTCAGCGAGGTGCATCTCCAAATGTGCTCACTGTTTCAGAAAGCCCTATGCTGCTCCGG 17315
Qy 2982 GCCACTTTACACTTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3041
Db 17316 GCCACTTTACACTTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 17375
Qy 3042 TGGTGGATGTCAGT 3055
Db 17376 TGCATGAGGGCGT 17389

RESULT 5

US-08-731-499-9
; Sequence 9, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10365 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: 1..10365
; LOCATION: 1..10365
; OTHER INFORMATION: /note= "Genomic Sequence Encoding
; OTHER INFORMATION: ZABCI"
; US-08-731-499-9

Query Match 49.0%; Score 1559.6; DB 8; Length 10365;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCAGAGAATCTCT 1541
Db 8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCAGAGAATCTCT 8345
Qy 1542 GAGGTATCACTTGCAGAGACATCAAGAGAAAAACCAACCGATGTTGCTGCTGAAGTCAA 1601
Db 8346 GAGGTATCACTTGCAGAGACATCAAGAGAAAAACCAACCGATGTTGCTGCTGAAGTCAA 8405
Qy 1602 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGCACAGTGGCAAC 1661
Db 8406 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGCACAGTGGCAAC 8465
Qy 1662 CAAAAATTTGAAAAAGATTTTGTGATGTGCAAAAGATGTTACAGCGAGTCCACCTGCAAA 1721
Db 8466 CAAAAATTTGAAAAAGATTTTGTGATGTGCAAAAGATGTTACAGCGAGTCCACCTGCAAA 8525
Qy 1722 GCAGCTTAAGGAGATGCTCTGTTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACC 1781
Db 8526 GCAGCTTAAGGAGATGCTCTGTTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACC 8585
Qy 1782 AGCACACAAAGATACTCAGGATTTCCATAAAAAATGCAGCTGATCAGCTGCTGATAAGT 1841
Db 8586 AGCACACAAAGATACTCAGGATTTCCATAAAAAATGCAGCTGATCAGCTGCTGATAAGT 8645
Qy 1842 GAATAAAAAACCTTACCCTGCTTACCTGGAGCTGTTAAAAAAGAGATCAGCAGTTGAAAC 1901
Db 8646 GAATAAAAAACCTTACCCTGCTTACCTGGAGCTGTTAAAAAAGAGATCAGCAGTTGAAAC 8705
Qy 1902 TCAGGCAATTAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCCGAGTGGCAG 1961
Db 8706 TCAGGCAATTAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCCGAGTGGCAG 8765
Qy 1962 TACCACCCATAAACCCTTGAAGTTAGCCCCCAAGAGAAAGCAAAACGGAGACCGCAGCTGACTG 2021
Db 8766 TACCACCCATAAACCCTTGAAGTTAGCCCCCAAGAGAAAGCAAAACGGAGACCGCAGCTGACTG 8825
Qy 2022 CAGATACAGGCCAAAGTGTGGATTGTACGAAAAACCTTTTAAATTTATCCGTGGGGGCTCT 2081
Db 8826 CAGATACAGGCCAAAGTGTGGATTGTACGAAAAACCTTTTAAATTTATCCGTGGGGGCTCT 8885
Qy 2082 TCACAAATGCCCCGCAATTTCTTTGAGTAAAGTTTGATCCCAAGTATCACTGTCCATT 2141
Db 8886 TCACAAATGCCCCGCAATTTCTTTGAGTAAAGTTTGATCCCAAGTATCACTGTCCATT 8945
Qy 2142 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACCCAGAGACTGGAGCATAA 2201
Db 8946 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACCCAGAGACTGGAGCATAA 9005
Qy 2202 ATACAAATCCTGACGTTTCATAAAAACTGTGCAAAACAAAGTCCCTGCTTAGAAGTGCAGTAC 2261
Db 9006 ATACAAATCCTGACGTTTCATAAAAACTGTGCAAAACAAAGTCCCTGCTTAGAAGTGCAGTAC 9065
Qy 2262 CGGATGCCCGCAGGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTCTCTGCTGCTGCTGCT 2321
Db 9066 CGGATGCCCGCAGGTTGCTGGGAAAGATGTGCTCCCTCCCTCTCTAGTCTCTGCTGCTGCT 9125
Qy 2322 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCGCATCTGCGAAGGGAAGCA 2381
Db 9126 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCGCATCTGCGAAGGGAAGCA 9185
Qy 2382 GAGCCCTCTCTGGGCCAGGCAAGGGCCCTCTGACTTCAGGGATAGACTCTAGCATTGAGC 2441

Db 9186 GAGCCCTCTCTGGCCAGGCAAGCCCTCTGACTTCAGGATAGACTCTAGCACTTTAGC 9245
Qy 2442 CCCAAGTAACCTGAAGTCCACAGACCAAGCAAGTAATGTGGGGTCCAAAGGGCCGCCAC 2501
Db 9246 CCCAAGTAACCTGAAGTCCACAGACCAAGCAAGTAATGTGGGGTCCAAAGGGCCGCCAC 9305
Qy 2502 CAGGCAACAGCAATCTGAGATGTTTCTTAACACCAAGTGTTCCTCCCTGACCGGATAAGAC 2561
Db 9306 CAGGCAACAGCAATCTGAGATGTTTCTTAACACCAAGTGTTCCTCCCTGACCGGATAAGAC 9365
Qy 2562 AAAAGAGCCCGAGACAAAATTAAGAACCTCTTCCAGTAGTCTCTTCCAGCCCAACCTCGG 2621
Db 9366 AAAAGAGCCCGAGACAAAATTAAGAACCTCTTCCAGTAGTCTCTTCCAGCCCAACCTCGG 9425
Qy 2622 CAGCAGTAACATCAATGTGTTCCATCGACTACCCCGCCCAAGACGACGCGGTGGGCACC 2681
Db 9426 CAGCAGTAACATCAATGTGTTCCATCGACTACCCCGCCCAAGACGACGCGGTGGGCACC 9485
Qy 2682 TCCGGGAGAGACTATTTCTGTAATCGGAGTGCAGCAATACTGACGAGCAATTTGGTGA 2741
Db 9486 TCCGGGAGAGACTATTTCTGTAATCGGAGTGCAGCAATACTGACGAGCAATTTGGTGA 9545
Qy 2742 GCCCCTTCCAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAAGCCCGGGC 2801
Db 9546 GCCCCTTCCAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTTGACCAAGCCCGGGC 9605
Qy 2802 CAATTACAGAGAGGCTATGACTTCCCAAGTACCATATGTTGACAGCATCATCACT 2861
Db 9606 CAATTACAGAGAGGCTATGACTTCCCAAGTACCATATGTTGACAGCATCATCACT 9665
Qy 2862 GTTACCGCAGGACTGTGTATCCCTCCAGCGGTGCTCCCAACCAAGTTCCTGAG 2921
Db 9666 GTTACCGCAGGACTGTGTATCCCTCCAGCGGTGCTCCCAACCAAGTTCCTGAG 9725
Qy 2922 CTCACGAGGTCGATTTCCAAATGTCGACTGTTCCAGAGCCCTATGTTGGTCCCG 2981
Db 9726 CTCACGAGGTCGATTTCCAAATGTCGACTGTTCCAGAGCCCTATGTTGGTCCCG 9785
Qy 2982 GCCACTTTACACTTGTGTCCTGCTGGTAGTCCAGCATCCAGCTCGAGTGTAGAGGTCT 3041
Db 9786 GCCACTTTACACTTGTGTCCTGCTGGTAGTCCAGCATCCAGCTCGAGTGTAGAGGTAT 9845
Qy 3042 TGGTGGATGTAGT 3055
Db 9846 TGCATGAGGGCGT 9859

RESULT 6

US-10-087-192-1706
; Sequence 1706, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIORITY FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIORITY FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIORITY FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1706
; LENGTH: 3016
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-1706

Query Match 47.4%; Score 1511.2; DB 13; Length 3016;

Best Local Similarity 71.5%; Pred. No. 0;
Matches 2194; Conservative 6; Mismatches 759; Indels 111; Gaps 12;
Qy 25 ATGCCAATCAATCCCTCTTAATGTACATGATGGCCAGAGTGAATTTGGCAGCTCTCTT 84
Db 1 ATGCCGACCCAGTCCCTCTCGTGTACATGACGCGGCCGGAAGTCTCTCAGCAGCTCTCTA 60
Qy 85 GGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATGAAGGACCGCTGTGTTCATTC 144
Db 61 GGCCTCCAGATGGAGGTTGATGCTGTGCCCATATAAAGGCGCGGTGGCAGTCCCTTC 120
Qy 145 CGAGCTACACAAGAAAAAATGTTCATCCAAATCGAGGGGTATATCCCTTTGATTTGCATG 204
Db 121 CGAGCTGTCTAGGAGAGAGCATGGCCGTGGCAGAGGGCCCATGCCCTCTGATTTGCATG 180
Qy 205 TTCTGCGCCAGACCTTCCACACATTTCAAGAAGCCTTAATAAATGTCTTATATGAACAC 264
Db 181 TTCTGCGCCAGGTTCTTCTCTCAGCGGAGGATCTCAGTCAGCAGCTGTCTGTCGAGCAC 240
Qy 265 CGGCTACCTCTGTGAACACGAGCTTCTTGGGTTGAAGCAGATATCTCAGTCCGGTT 324
Db 241 CGGCTCACCTCTGCGAGCCAGCTGTCTGCGTGTGGAGCGCGAGTACCTTAAGTCCCTT 300
Qy 325 GATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAAAATTCGAAGA --- AATGAA 381
Db 301 GATAAAGTCTGGAGCCACAGAGCCAGCATTTGGAGAAAGATGGCGAAGACCCCGAGGAG 360
Qy 382 TTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATG 441
Db 361 TTGAGCTGTGATGTGTGGCGCAGACATTTCCAGTGGCTTTTGTATGTTGAGAGCCACATG 420
Qy 442 AGAACACAAAGATTTTTCATCTTACCTTACCGGTGTAAATGTCGGGAAGAGMTTSSRSAG 501
Db 421 AAGAAGCATAAAGGACTCTCTTCAAGTATGGGTGACAGCATGTGCGGAGGAGATTTCAAGAG 480
Qy 502 CTTGGTTTCTTAAAAATCACATGCGGACACATATGCAAAATCGGGGSCCAGAGAGCAA 561
Db 481 CCGTGGTTCTTGAAGAACCATATCGGACACACATATGGCAAGTCTGGCACACAGGAGCAG 540
Qy 562 CTGACGAAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGACCGG 621
Db 541 CTTCAGCAAGCATGAGG --- AGTCCAGTCCACATCAATGAAGTGTCCAGCGCACGCC 597
Qy 622 GCGAGAGCATCTCTCTCTTACAAATCTGATGTTTGTGGCTTCTTATTTCCAAAT 681
Db 598 CTTGGAGCATCTCCAGCCCTTACAGATCTGATGTTCTGCGGCTTCTCTTCTCCAAAT 657
Qy 682 AAAGAAAGTCTAATTTGAGCACCGCAAGGTGCACACCAAAATAAATGCTTTTCGTACACG 741
Db 658 AAGCAGAGCTCTATTTGAGCACAGCAAGGTTTACGCCAAAGAAACTGTCCCGAGTCCGAGC 717
Qy 742 AGCGCGCAGACAGACTCTTCCACAAGGAGGAATGCGCTCTCCAGGAGGAGACTTCTTCGAG 801
Db 718 AACGTTGCTGCTGATGACCCAGAGAGAAACCCAGCTCCCGAGGGAAGAGTGTCTGCAG 777
Qy 802 TTGTTCAACTTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCTGTTCAGATGATC 861
Db 778 TTTTGAACCTTGAGACCCAGATCAACTGACAGGTAGTACAGTGAAGCCCATGACCTGCATA 837
Qy 862 CCTCAGCTCGATCCGTTTCAACCACTTCCAGCTTGGCAGCTGGCTTACCAAGGAAAGATT 921
Db 838 CCTCAGCTTGACCCGTTTCCACCCTTACAGGATGGCAGTGGCTTACCAAGGAAAGATTG 897
Qy 922 GCCATTTTGC --- AAGAAAGTGAAGGAATTCGGGGAAGAGGAGCACCGACCAACAGCAT 978
Db 898 GCGTTGCCAGGAAGAGGTGAAGAGTCAAGGCAAGAGGAGGAGCACAGACATGACGAC 957
Qy 979 TCGAGTTCCGAGAGGAGCTTGGAG --- TAGGGTGGGGAAATATGTTGGGGTAAGCGGAGAGTGC 1003
Db 958 TCATGCTCAGAGAAAGAGGAACTAGGGGAAATATGTTGGGGTAAGCGGAGAGAGTGC 1017
Qy 1004 -----AAACAATAGGGCAGTTGTCCAGGCTCTCCAGAGAAAGAGAGTGC 1053

Db 1018 GGAAGTCAAAACAAGTAAAGCAGTTGTCCAGGTCTCTCCCAAGACAAAGAGAGCCT 1077
Qy 1054 AAACACTCCACGGCGAGCGCCCTCGTGTGACCGGATCCCAAGTTACCAGTAGCAAG 1113
Db 1078 AGACATGCTAATAGTAGAGTCCCTTCTGGGATAGTAGACCCCAAGTTGTCCAGTAGCAAG 1137
Qy 1114 GAGAAAGCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACTCAACACGACTGGTC 1173
Db 1138 GAGAAAGCCACGACTGTTCTGAGTGCAGCAAGGCTTCAGACATACCAACGACTGGTC 1197
Qy 1174 TTGCACTCCAGGGTCCACAAGAGGACGGAGGCGCGGCGAGTGGCCACCATGTCT 1233
Db 1198 CTGCACCTCGAGGTGCACAGGAAGGACAGGAGGACTGATGCCCTGTCAACCACCATGGCT 1257
Qy 1234 GTGACGGGAGGACCGGGAGCGTGTCTCTGACCTCGCGCGCCCTCTGGATGAAAT 1293
Db 1258 GTGATGCAAGGACGCTGGACCTGTCTCCACGACCTTCAGCACTCTGGAAGACAGT 1317
Qy 1294 GGAGCGGTGGATCGAGGGGAAGGTGTTCTGAAGACGGATCTGAGGATGGGCTTCCCGAA 1353
Db 1318 GGGCCGGGACCGA--GAAGGGGCTCTGAAGCGGTCTGAGGATGGACTCCCTGAC 1374
Qy 1354 GGAATCCATCTGGATAAATATGATGAGAGGAAATAAACAATCTTACATCTTCAGA 1413
Db 1375 GGGCTCCATTTGGATAAATATGATGAGAGGAAAGCGAAGCCCTCCCTCCCTCGAGA 1434
Qy 1414 GAGTGTAGTTATTGTGGAAAGTTTTCCGTTCAAAATTTATACCTCAATATCTCTAGA 1473
Db 1435 GAGTGTAGTTACTGTGGCAAGTTTTCGTTCAAACTATTTACCTCAATATCTCTAGA 1494
Qy 1474 ACGCATACAGGTGAATAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCAGAG 1533
Db 1495 ACGCATACAGGTGAATAACCATACAAATGTGAATTTCTGTGATGTCGCGAGCCAGAG 1554
Qy 1534 ACATCTCTGAGGTATCACTTGGAGAGACATCAAGAGGAAACAAAC---CGATGTTGCT 1590
Db 1555 ACATCTCTGAGGTACCACTTGGAGAGACATCAAAAGACAGAGCGCGGTGGATGCTGCC 1614
Qy 1591 GCTGAAGTCAAGACGATGTTAAATACAGACACTGAAGTGCACATTTAA---CCGCT 1647
Db 1615 GCTGAGTCCAAAGTGAAGCGCGGAGCCAGAGCGCGAGGATGCGCTACTAACCGGTGCT 1674
Qy 1648 GACAGTGGCCAAACCAAAATTTGAAAGATTTTGTGATGTCGCAAGATGTTACAGGC 1707
Db 1675 GACAGTGGCCAGACCAAAATTTAAGAGATTTCTGTGATGTCGCAAGATGTTAAGGA 1734
Qy 1708 AGTCCACTGCAAGACAGCTTAAGAGATGCTTCTGTTTTTCAGATGTTCTGGGCAGC 1767
Db 1735 AGCCACCTGCCAAGCAGCTTAAGGAGATGCTTCTGTTCCAGAGTGTTC----- 1786
Qy 1768 GCTGCTCTCACGAGCACAAAGATACTCAGGATTTCCATTAATAATGACGCTGATGAC 1827
Db 1787 ----TCTCACCAGCACACAGCAACGATACTCAGGATTTCCATAACATGACGCTGAT--- 1839
Qy 1828 AGTGTGATAAAGTGAATAAAACCCCTACCCCTGCTTACCTGGACCTCTTAAAAAAGAGA 1887
Db 1840 AGTGTGATAAAGCGAGAGAGCCCTGCTGCTTACTTATCTGGACATCGAGAGAGA-- 1897
Qy 1888 TCAGAGTTGAAACTCAGGCAAAATTAACCTCATCTGTAGAAACAAGCGGATGTTACTCCT 1947
Db 1898 -AAGCAGGGAGCCTCAGGCCAGCAGCCCTGTCTGCAGACTAGAGGGGTTGGGTCTCTTA 1956
Qy 1948 CCTCCGATGGCAGTACCACCATACCTTGAAGTTAGCCCCCAAGAGAGCAACGAG 2007
Db 1957 GCACGGGAAGCTGG-----CCATAGGGGAGAGATGGAT 1989
Qy 2008 ACCGAGCTGACTGCAGATACAGGCCAAGTGTGGATTTGTACGAAAAACCTTTTAAATTTA 2067
Db 1990 CAGGATGCTGACTACAGACATTAAGCCCGTGTGACTGCCAGGACAGCCCTTTGATCTA 2049
Qy 2068 TCCGTGGGGGCTTTTCAAAATTTGCCCGCAATTTCTTTGAGTAAAGTTTGAATCCAAGT 2127
Db 2050 TCCCTTGGGCGGCTCCAGCGCTGTCTCTGCAATCTCTTTGAGCAAGTGTCTGATCCCGAGC 2109

Qy 2128 ATCACTGTCTCAATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTAAATGATGCACAG 2187
Db 2110 ATTGCTTGCCTTCTTTGTACTTTCAAGACCTTTTATCCGNAAGTCTTATGATGCACAG 2169
Qy 2188 AGACTGGAGCATAAATACAACTCTGACGTTTCATAAAACCTGTGAAAACAAGTCTTGCTT 2247
Db 2170 AGACTTGAGCACAGGTACAAACCTTGACCCGACACAAGAACGCGCAGCAGCAAGTCTGTGCTG 2229
Qy 2248 AGAAGTCACTACCGGATGCCCGCAGCCGTTGCTGGGAAAGATGTCCTCCCTCTCT 2307
Db 2230 AGGAACAGGGGTACCGGGTGCCTCCGGCTTTGCTGGGAAAGATGTCCTCTCTCTGCT 2289
Qy 2308 AGTTTCTGTAAACCAAGCCCAAGCTCTGCTTTTCCCGCGCAGTCCAAATCCCTGCCATCT 2367
Db 2290 GGCTGCACAAGCCCAAGCCAGACTGCTCTTCTACCACTCAAGTCCCTGCACCTCA 2349
Qy 2368 GCGAAGGGGAAGCAGAGCCCTCTTGGCCAGGCAAGGCCCTCTGACTTCAGGAGATAGAC 2427
Db 2350 GAGAAGGCTCGGCAGGGGGCTCGGGGCCAAGCAAGCAAGCAGACTTCAGGACCCAGAC 2409
Qy 2428 TCTAGCACTTTAGCCCAAGTAACTGAAGTCCACAGACACAGACAGAGATGTGGGGTC 2487
Db 2410 AACAGCACTTTAGCCCCCAAGTAACCTGAAGTCAACAGGTCAACACCCAAATGCTGGGGC 2469
Qy 2488 CAAGGGCGCGCCACCGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCT 2547
Db 2470 ACCAG---TGCCACACGCGCAGCAGCTCAGAGTTGTTTCCCAAGGTGGCGTCCCTGCT 2526
Qy 2548 GCACCGGATTAAGCAAAAAGACCCGAGACAAAATTTGAAAACCTCTTCCAGTAGCTCTTCT 2607
Db 2527 GCTATGATTAAGTGAAGAGACCTGAGCCAAAACCTGAAGTCCCTTACCAGCCTCCCGCT 2586
Qy 2608 CAGCCCACTCTCGCAGCAGCTAACTCAATGGTTTCCATCGACTACCCCGCCCAAGACAC 2667
Db 2587 CAGTCCCTCTCAGCAGCTAATAATAGCAACGGTTCTGTGTAGTATCCCGTGAAGTTGAC 2646
Qy 2668 AGCCCGTGGGCACCTCCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATACTGCA 2727
Db 2647 GGCCCATGGGCACAGCAAGGAGAGACTACTACTGCCATCGGAATTTCTGGCAGTGGCGCA 2706
Qy 2728 GCAGAAATTTGTTGAGCCCTTTCCAAAAGACTGAAGTCCAGCGTGGTGGCCCTTGACGTT 2787
Db 2707 GCAGAGTACAGTGAAGCACATCCCAAAAGACTCAAGTCCAGTGGCGGTGCTCCCTGACACA 2766
Qy 2788 GACCAGCCCGGGCCAAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGA 2847
Db 2767 GAGCATGCAGGACCAATGGCAGAAAGGGCTTTGAGCTCCCCAAGTACCATGTGTGTCAGG 2826
Qy 2848 GGCAATCACATCACTGTTTACCGCAGGACTGTGTATCCGTCCGAGGGCGCTGCCTCCCAAA 2907
Db 2827 AGCATCACTCTTGTCTACCAACAGAGTGTGTGGCCACCGCTGTGCTGCCCCACAAA 2886
Qy 2908 CCAAGGTTCTGAGCTCCAGCAGGTGATTTCTCAAAATGTCGATGTTTCAGAGGCC 2967
Db 2887 GCCCGTTTCTGAGCCCTTGGGAGGTGAGTCAAGTGTGTGGCTGTGTCAGAGGCC 2946
Qy 2968 TATGTTGCTCCGGCCCACTTTTACACTTGTGTGCTGCTGTGTAGTCCAGACTCCAGCTCG 3027
Db 2947 TACAGTGCCTCTGAGCCCTGTATACCTGTGTGACCCGTTGGGACACGCGAGGAGGCCCA 3006
Qy 3028 ACGTTAGAAG 3037
Db 3007 GCCCTTGAAG 3016

RESULT 7

US-08-731-499-3
; Sequence 3, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin

;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: US 09/798,586
;; PRIOR FILING DATE: 2001-03-02
;; NUMBER OF SEQ ID NOS: 2059
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1705
;; LENGTH: 26345
;; TYPE: DNA
;; ORGANISM: Mus musculus
US-10-087-192-1705

Query Match 23.0%; Score 733; DB 13; Length 26345;
Best Local Similarity 70.2%; Pred. No. 7.5e-206;
Matches 1106; Conservative 0; Mismatches 415; Indels 54; Gaps 7;

Qy 1482 AGGTGMAAAACCATCAATCTGAAATTTTGTGAAATATGCTGCAGCCAGAGAGACATCTCT 1541
Db 14832 AGGTGMAAAACCATCAATCTGAAATTTTGTGAAATATGCTGCAGCCAGAGAGACATCTCT 14891

Qy 1542 GAGGTATCACTTGGAGAGACATCAAGAGAAAAACAAAC---CGATGTTGCTGTGAAGT 1598
Db 14892 GAGGTACACCTTGGAGAGACATCAAGAGACAGCAGCCGGTGGATGTCGCGCTGAGTC 14951

Qy 1599 CAAGAACGATGGTAAAAATAGGACACTGAAGATGCATATTA---CCGCTGACAGTGC 1655
Db 14952 CAAAAGTGAAGCCGAGCCAGGAGCCGCGCTACTAAACGGCTGCTGACAGTGC 15011

Qy 1656 GCAACCAAAATTTGAAAGATTTTGTGATGTTGCCAAGATGTTACAGGCAGTCCAC 1715
Db 15012 GCAGACCAAAATTTTAAAGAGATTTCTTGAATGTTGCCAAGATGTTAAGGAAAGCCAC 15071

Qy 1716 TGCAAGCAGCTTAAGAGAGATGCTTCTGTTTTTTCAGAAATGTTCTGGGCAGCGTGTCT 1775
Db 15072 TGCCNAGCAGCTTAAGAGAGATGCTTCTGTTTCCAGATGTTCTC-----TCTC 15119

Qy 1776 CTCACACACACAAAGATATCTAGGATTTCCATAAAAATGACGCTGATGACAGTCTGA 1835
Db 15120 ACCAGCACACAGCAACGATCTCAGGATTTCCATAAACATGACAGCTGAT---AGTGTCTGA 15176

Qy 1836 TAAAGTGAATAAAACCTTACCCCTGCTTACTGACCTGTTTAAAGAGATCAGCAGT 1895
Db 15177 GAAAGCGAGGAAGAGCCCTGCCCCCTACTTATCTGACATGACAGAAAGAAA---GCGAG 15233

Qy 1896 TGAAGCTCAGGCAATAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCTCCCGA 1955
Db 15234 GGAGCTCAGCCAGCAGCCCTGCTGACACTAGAGGGGTTGGTCTTTAGCAGCGGA 15293

Qy 1956 TGGCAGTACCACCCATAACCTTGAAGTTAGCCCCAAAGAGAGAAACGAGACCGCAGC 2015
Db 15294 AGCTGG-----CCATAGGGAGAGATGGATCAGGATGC 15326

Qy 2016 TGACTGCAGATACAGGCCAAGTGTGGATTTGTACAGAAAAACCTTTAAATTTATCCGTGG 2075
Db 15327 TGACTACAGACATAAGCCCGGTGCTGACTGCCAGACAGGCTTTGAAATCTATCCCTTGG 15386

Qy 2076 GCCTCTTCACAATTCGCCGCAATTTCTTTAGTAAAGTTGATTCCAAAGTATCACCTG 2135
Db 15387 GCCGCTCAGCGCTGCTGCAATCTCTTTGAGCAAGTGTCTGATCCCCAGCATTGCCTG 15446

Qy 2136 TCCATTTTGTACCTTCAAGACATTTTATCCAGAAATTTTAAATGATGACACAGAGACTGA 2195
Db 15447 CCCCCTTTTGTACTTTTCAAGACCTTTTATCCGGAAGTCTTATGATGACACAGAGACTTGA 15506

Qy 2196 GCATAAATACATCTCTGAGTTCATAAAACTGTGAAACAGTCTTGTCTAGAGTGC 2255
Db 15507 GCACAGGTACAAACCTTGACCCGCCAAGAACGGCAGCAGCAAGTCTGTGCTGAGAAACAG 15566

Qy 2256 AGGTACCCGATGCCGCCAGCGCTTGTCTGGAAAGATGTCCTCCCTCTCTAGTTTCTG 2315
Db 15567 GCGTACCCGGTGCCTTCCGGCTTTGCTGGAAAGATGTCCTCTCTGTCTGGCTGCA 15626

Qy 2316 TAAACCCAGGCCAAGTGTCTGTTTTCCCGCGCAGTGCATAATCCCTGCAATCTGCAAGGG 2375
Db 15626 TAAACCCAGGCCAAGTGTCTGTTTTCCCGCGCAGTGCATAATCCCTGCAATCTGCAAGGG 2375

Db 15627 CAAGCCCAAGCCCAAGACTGCCTTTCTCACACACTCGAAGTCTCTGCACCTCAGAGAAGGC 15686
Qy 2376 GAAGCAGAGCCCTCTCTGGGCCAGCAAGGCCCTCTGACTTTACGGGATAGACTCTTAGCAC 2435
Db 15687 TCGCAGGGGGCTCTGGGGCCAGCAAGCACCACGACTTCAGGACACAGACAGCAC 15746
Qy 2436 TTTAGCCCCAAGTAACTGAAGTCCCAAGACCAACAGAGAAATGTGGGGTCTCAAGGGGC 2495
Db 15747 TTTAGCCCCAAGTAACTGAAGTCAACAGGTCAACACCCCAATGCTGGGGCCAGCAG--- 15803
Qy 2496 CGCCACCCAGCCCAAGCAATCTGAGATGTTTCTTAAACACAGTGTTCCTCCCTGCACCGGA 2555
Db 15804 TGCCACCCAGCCAGCAGCTGAGATGTTTCTTCCAAAGGTGGCTTCCCTGCTGTATGGA 15863
Qy 2556 TAAGACAAAAGAGCCCGAGACAAAATTAACACCTCTTCCAGTAGTCTCTTCTCAGCCCCAC 2615
Db 15864 TAAAGTGAAGAGACTGAGCCAAAACCTGAAGTCCCTACAGCCTCCCGTCTCAGTCCCC 15923
Qy 2616 CCTCGGACGAGTAAACATCAATGTTCCATCGACTACCCCGCCCAAGAACGACAGCCGTG 2675
Db 15924 CCTCAGCAGTAAATATAGCAACGGTCTGTTGAGTATCCCGTGAAGGTTGACGGCCCATG 15983
Qy 2676 GGCACCTCCGGGAAGAGACTATTTCTGTAATCGAGTCCAGCAATACTGCAGCAGAAAT 2735
Db 15984 GGCACAGCAAGGGAGAGACTACTACTGCTCATCGGAATTTCTGGCAGTGGCCGAGCAGAGTA 16043
Qy 2736 TGGTGAGCCCTTCCAAAAGACTGAAGTCCAGCGTGGTTCCTTGAAGTTCGACGTTGACCCAGCC 2795
Db 16044 CAGTGAGCCACATCCCAAGAGACTCAAGTCAAGTGGGTGCTTCTTGGACACAGAGCATGC 16103
Qy 2796 CGGGGCCAAATPACAGAGAGAGCTATGACCTTTCCAAAGTACCATATGTTCAAGGAGCATCAC 2855
Db 16104 AGGCACCAATGGCAGAGGGGCTTTGAGCTCCCCAAGTACCATGTGTTGGTCAAGAGCATCAC 16163
Qy 2856 ATCACTGTTACCGAGCAGCTGTGTATCCGTCCAGCGCTGCTCCCAACCAAGGTT 2915
Db 16164 CTCTTGTCTACCAACAGAGTGTGTGCGCCACCGCTGTGCTGCTGCCCAAGCCCGTTT 16223
Qy 2916 CCTGAGCTCCAGCAGGTCGATTTCTCCAAATGTGTGCTGACTGTTTCAAGAGCCCTATGTTGG 2975
Db 16224 CTTGAGCCCTGGGAGGTGGAGTCAACAGTGTGTGTTGGCTGTGAGAGCCCTACAGTGC 16283
Qy 2976 CTCCGGGCCAATTTACACTTGTGTGCTGTGCTGTGTTGTTGAGTCCAGCATCCAGTCCAGT 3035
Db 16284 CTCTGGACCCCTGTATACCTGTGGACCCGTGGGACACGACAGGAGCGACCCAGCCCTTGA 16343
Qy 3036 AGTCTTGGTGGATG 3050
Db 16344 AGGTATTACTGTAGG 16358

RESULT 9

US-10-040-739-520
; Sequence 520, Application US/10040739
; Publication No. US20020173635A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoY, John

LaVallie, Edward

Racie, Lisa

Merberg, David

Treacy, Maurice

Spaulding, Vikki

TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS

NUMBER OF SEQUENCES: 1519

CORRESPONDENCE ADDRESS:

ADDRESS: Genetics Institute, Inc.

STREET: 87 CambridgePark Drive

CITY: Cambridge

STATE: Massachusetts

COUNTRY: U.S.A

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/040,739
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/036,520
FILING DATE: 03-JUN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 520:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 520:
US-10-040-739-520

Query Match 13.9%; Score 444.4; DB 13; Length 469;
Best Local Similarity 99.8%; Pred. No. 5.4e-121;
Matches 445; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1032 CTCGCAAGAAAGAGAGTGCACACCTCCACGCGGAGCGCCCTCGGTGGAGCGGA 1091
DB 19 CTAGCAAGAAAGAGAGTGCACACCTCCACGCGGAGCGCCCTCGGTGGAGCGGA 78
QY 1092 TCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTT 1151
DB 79 TCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTT 138
QY 1152 CAGAACCTTACCAACAGTGGTCTTGCACTCCAGGGTCCACAGAGGACCGGAGGCGCG 1211
DB 139 CAGAACCTTACCAACAGTGGTCTTGCACTCCAGGGTCCACAGAGGACCGGAGGCGCG 198
QY 1212 CGCGGAGTGGCCACCATGTCTGTGGAGGAGGAGCGCGGAGCGTCTCTCTGACCT 1271
DB 199 CGCGGAGTGGCCACCATGTCTGTGGAGGAGGAGCGCGGAGCGTCTCTCTGACCT 258
QY 1272 CGCGCCCTCTGGATGAAATGGAGCCGTGGATCGAGGGAAGGTGTTCTGAAGACGG 1331
DB 259 CGCGCCCTCTGGATGAAATGGAGCCGTGGATCGAGGGAAGGTGTTCTGAAGACGG 318
QY 1332 ATCTGAGATGGCTTCCGAGGAATCCATCTGGATGAAATGATGATGAGGAGGAAAT 1391
DB 319 ATCTGAGATGGCTTCCGAGGAATCCATCTGGATGAAATGATGATGAGGAGGAAAT 378
QY 1392 AARACATCTACATCTCAAGAGAGTGTAGTTATGTGGAAGTCTTCCGTTCAAAATTA 1451
DB 379 AARACATCTACATCTCAAGAGAGTGTAGTTATGTGGAAGTCTTCCGTTCAAAATTA 438
QY 1452 TTACCTCAATATTCACTCCAGAACGC 1477
DB 439 TTACCTCAATATTCACTCCAGAACGC 464

RESULT 10
US-10-029-386-9711/c
; Sequence 9711, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23411
; LENGTH: 251

; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9711
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: SWISSPROT HIT: O75362, EVALUE 2.00e-67
; OTHER INFORMATION: NT HIT: g11421959, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF237807.1, EVALUE 0.00e+00
US-10-029-386-9711

Query Match 12.6%; Score 401.4; DB 16; Length 530;
Best Local Similarity 99.8%; Pred. No. 3.7e-108;
Matches 402; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 968 ACAACGACGATTCGAGTTCGAGAGGAGCTTGGAGAAACAAATAAGGCGAGTTGTGCAG 1027
DB 530 ACAACGACGATTCGAGTTCGAGAGGAGCTTGGAGAAACAAATAAGGCGAGTTGTGCAG 471
QY 1028 GCCTCTCGCAGAGAAAGAGTGCACAACTCCACGCGGAGCGCCCTCCGTGGAGC 1087
DB 470 GCCTCTCGCAGAGAAAGAGTGCACAACTCCACGCGGAGCGCCCTCCGTGGAGC 411
QY 1088 CGGATCCCAAGTTACCCAGTAGCAAGAGGAGAGCCCACTCACTGCTCCGAGTCCGCGCAAG 1147
DB 410 CGGATCCCAAGTTACCCAGTAGCAAGAGGAGAGCCCACTCACTGCTCCGAGTCCGCGCAAG 351
QY 1148 TTTTCAGAACCTACCAACGAGTGGTCTTGTGACTCCAGGGTCCACAAGAGGACCGGAGGG 1207
DB 350 TTTTCAGAACCTACCAACGAGTGGTCTTGTGACTCCAGGGTCCACAAGAGGACCGGAGGG 291
QY 1208 CGGCGCGGAGTCCGCCACCATGTCTGTGGAGGAGGAGCGCGGGAGCGTGTCTCTCTG 1267
DB 290 CGGCGCGGAGTCCGCCACCATGTCTGTGGAGGAGGAGCGCGGGAGCGTGTCTCTCTG 231
QY 1268 ACCTCGCGCCCTCTTGATGAAATGAGCCGTGGATCGAGGGAAGGTGGTCTCTGAAG 1327
DB 230 ACCTCGCGCCCTCTTGATGAAATGAGCCGTGGATCGAGGGAAGGTGGTCTCTGAAG 171
QY 1328 ACGATCTGAGATGGGCTTCCGAGGAATCCCATCTCGATTA 1370
DB 170 ACGATCTGAGATGGGCTTCCGAGGAATCCCATCTCGATTA 128

RESULT 11
US-10-029-386-23411/c
; Sequence 23411, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23411
; LENGTH: 251


```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EST_HUMAN HIT: W05407.1, EVALUATE 1.00e-117
; OTHER INFORMATION: NT HIT: g11421959, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: O75362, EVALUATE 1.00e-38
US-10-029-386-23411

Query Match          7.9%; Score 251; DB 16; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.5e-63;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1093 CCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTC 1152
Db 251 CCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTC 192

Qy 1153 AGAACTTACCCAGTGGTCTTGACATCCAGGGTCCACAGAAAGACCCGAGGCGCGC 1212
Db 191 AGAACTTACCCAGTGGTCTTGACATCCAGGGTCCACAGAAAGACCCGAGGCGCGC 132

Qy 1213 CGGGAGTCCGCCACCATCTCTGTGACGGGAGGCGAGCGGTGTCTCTGACCTC 1272
Db 131 CGGGAGTCCGCCACCATCTCTGTGACGGGAGGCGAGCGGTGTCTCTGACCTC 72

Qy 1273 GCGGCGCTCTGGATGAAATGGAGCGTGGATCGAGGGGAAGGTGGTCTGAAAGCGGA 1332
Db 71 GCGGCGCTCTGGATGAAATGGAGCGTGGATCGAGGGGAAGGTGGTCTGAAAGCGGA 12

Qy 1333 TCTGAGGATGG 1343
Db 11 TCTGAGGATGG 1

RESULT 12
US-09-783-590-9057
; Sequence 9057, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9057
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (65)
```

```
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (75)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (103)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (108)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (113)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (193)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (204)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (212)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9057

Query Match          5.1%; Score 164; DB 9; Length 267;
Best Local Similarity 90.4%; Pred. No. 1.2e-37;
Matches 226; Conservative 0; Mismatches 18; Indels 6; Gaps 5;

Qy 2315 GTAAACCCAGCCCAAGTCTGCTTTCCCGGCGAGTCCAAATCCCTGCCATCTCGGAGG 2374
Db 18 GTAAACCCAGCCCAAGTCTGCTTTCCCGGCGCA-TCCAAATCCCTGCNATCTCGGAANG 76

Qy 2375 GG-AAGCAGAGCCCTCTCTGGGCCAGGCAAGCCCTCTGA--CTTCAGGGATAGACTCTA 2431
Db 77 GGAAGCAGAGCCCTCTCTGGGCCAGGAGGNCCTWTGAACCTTCAGGGAATAGACTCTA 136

Qy 2432 GCACCTTTAGCCCCAAGT-AACCTGAAGTCCACAGACACAGAGAAATGTGGGGTCCAA 2490
Db 137 GCACCTTTAGCCCCAAGTAAACCTGAAGTCCACAGACACAGAGAAATGTGGGGTNCAA 196

Qy 2491 GGGGCGGCCACCCAGGCAACAGCAATCTG-AGATGTTTCTTAAACCAGTGTTCCTCCCTGC 2549
Db 197 GGGGCGGCCACCCAGGNAACAGCAATTTGAAGATGTTTCTTAAACCAGTGTTCCTCCCTGC 256

Qy 2550 ACCGGATAAG 2559
Db 257 AACGGTTAG 266

RESULT 13
US-10-172-118-1511
; Sequence 1511, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Horigue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1511
; LENGTH: 6033
; TYPE: DNA
; ORGANISM: Homo sapiens
```



```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 93
; LENGTH: 8156
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-93

Query Match      3.2%; Score 102.4; DB 14; Length 8156;
Best Local Similarity 52.4%; Pred. No. 2.3e-18;
Matches 289; Conservative 5; Mismatches 246; Indels 12; Gaps 3;

Qy 379 GAATTTAGCTGTGAGGTATGTGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCAC 438
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 439 ATGAGACACACAAGATTCTTCTACTTACGGGTGTACATGTGCGGAGAGAGTTTSPRS 498
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1117 ATGAAGAAGCACCAGGGGCTCTTGACCAACGGCTGCCACATCTGCGGCCGTAGGTTCAAG 1176
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 499 SAGCCTTGGTTTCTTAAATAATCACATGCGGACACATATGCAAAATCGGGGCCCAGAAGC 558
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1177 GAGCCTTGTTCTTCAAGAACCAACATGAAGGGGCAAGGCCCAAGACGGGCGAGCAAGAAC 1236
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 559 AAATGCGACGAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGCAC 618
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1237 AGGCCCAAGAGTAGCTGGACCCCA---TCGCCACCATCAACAACGTTGCTCCAGGAGAG 1293
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 619 CGGCGCGAGACATCTCTCTCTTACAAATCTGCATGGTTTGTGGCTTCTATTTCAC 678
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1294 GTGATCGTCGCGGCTGAGCCTCTACGAGGTCTGCGCCAAAGTCGGGAACCTGTTTACA 1353
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 679 AATAAGAAAGTCTAATTTAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTGCGTACC 738
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1354 AACCTGGACAGCTTGAACGCCCAACATGCCATCCACCCGACAGTCGAGCC---CAGCCGC 1410
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 739 AGCAGCGCGCAGACAGACTCTCCACAGGAGGAATGCCGTCTCTGAGGGAGGACTTCTG 798
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1411 ACGCGCGCCCGCGGAGGAGGGGCGGAGGGGCCCTCGGACACCAAGCAGTCTTCTC 1470
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 799 CAGTTGTTCAACTTGAGACCAAAATCTCACCTGAAACG-----GGGAAGAAGCCTGTC 852
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1471 CAGTGCCCTGAACCTGAGGCCGTGCGCGCCCGCGGACTCGTCCCTGGCACGCGAGCCGGA 1530
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 853 AGATGCATCCCTCAGCTCGATCCGTTACACCTTCCAGGCTTGGCAGCTGGCTACCAAA 912
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 913 GGAAAAGTTGCC 924
Db ||| ||| ||| |||
Qy 1591 GGTAAGGTGGCC 1602
Db ||| ||| ||| |||
```

Search completed: July 4, 2005, 11:51:05
Job time : 1863.96 secs